NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

THESIS

THE DIGITAL LIBRARY PHENOMENON: OPPORTUNITIES AND IMPLICATIONS FOR THE NAVAL SERVICE

by

Robert E. Norris
and
David S. West

June 1996

Principal Advisor: Hemant K. Bhargava

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE June1996	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE **THE DIGITAL LIBRARY PHENOMENON: OPPORTUNITIES AND IMPLICATIONS FOR THE NAVAL SERVICE**	5. FUNDING NUMBERS
6. AUTHOR(S) Norris, Robert E., and West, David S.	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey CA 93943-5000	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

This thesis examines the emerging field of work encompassed by the term "Digital Library," and offers a plan for developing a Naval Service Digital Library. The amount of data and processing capabilities, available via networking technology, already defies description and continues to grow daily. As access to electronic resources and their diversity increase, a void in electronic Information Management principles and technologies has been uncovered. Participants in the global Digital Library (DL) movement are striving to adapt the principles of Library Science from locally controlled and accessed resources (books, magazines, videos, etc.) to remotely-shared electronic media and data processing systems. This thesis specifically addresses the movement's background, current initiatives and technologies (circa 1996).

The Naval Service can benefit immediately from monitoring and exploiting the DL technologies being developed world-wide. There are tremendous economies to be reaped in meeting our non-tactical, day-to-day information needs. To date, Navy and Marine Corps DL-related projects are narrowly focused by organization and limited to tactical, engineering and research information needs. By consolidating these efforts with a unifying vision and cooperative intent, a Naval Service Digital Library (NSDL) can be constructed. The NSDL would benefit all service members, in both their professional and personal lives, by providing a gateway to millions of resources that are compatible, searchable and ready for use. This thesis recommends an organizational structure and management strategy for developing a Naval Service Digital Library.

14. SUBJECT TERMS Digital Library, NSDL, DL			15. NUMBER OF PAGES 126
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18 298-102

Approved for public release; distribution is unlimited.

THE DIGITAL LIBRARY PHENOMENON: OPPORTUNITIES AND IMPLICATIONS FOR THE NAVAL SERVICE

Robert E. Norris
Commander, United States Navy
B.A., University of New Mexico, 1977

David S. West
Captain, United States Marine Corps
B.A., Virginia Military Institute, 1990

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
June 1996

Author: Robert E. Norris

David S. West

Approved by: Hemant K. Bhargava, Principal Advisor

George Zolla, Associate Advisor

Reuben T. Harris, Chairman
Department of Systems Management

ABSTRACT

This thesis examines the emerging field of work encompassed by the term "Digital Library," and offers a plan for developing a Naval Service Digital Library. The amount of data and processing capabilities, available via networking technology, already defies description and continues to grow daily. As access to electronic resources and their diversity increase, a void in electronic Information Management principles and technologies has been uncovered. Participants in the global Digital Library (DL) movement are striving to adapt the principles of Library Science from locally controlled and accessed resources (books, magazines, videos, etc.) to remotely-shared electronic media and data processing systems. This thesis specifically addresses the movement's background, current initiatives and technologies (circa 1996).

The Naval Service can benefit immediately from monitoring and exploiting the DL technologies being developed world-wide. There are tremendous economies to be reaped in meeting our non-tactical, day-to-day information needs. To date, Navy and Marine Corps DL-related projects are narrowly focused by organization and limited to tactical, engineering and research information needs. By consolidating these efforts with a unifying vision and cooperative intent, a Naval Service Digital Library (NSDL) can be constructed. The NSDL would benefit all service members, in both their professional and personal lives, by providing a gateway to millions of resources that are compatible, searchable and ready for use. This thesis recommends an organizational structure and management strategy for developing a Naval Service Digital Library.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS

ADL:	Alexandria Digital Library
ARPA:	Advanced Research Projects Agency
CALL:	Center for Army Lessons Learned
CDR:	Commander
CERES:	California Environmental Resource Evaluation System
CIA:	Central Intelligence Agency
CIIR:	Center for Intelligent Information Retrieval
CNN:	Cable New Network
CNRI:	Corporation for National Research Initiatives
CO:	Commanding Officer
CVN:	Carrier, Fixed Wing, Nuclear
DEC:	Digital Equipment Corporation
DIRECT:	Desktop Information Resources and Collaboration Technology
DISA:	Defense Information Systems Agency
DL:	Digital Library
DLI:	Digital Library Initiative
DoD:	Department of Defense
DSS:	Decision Support System
DTD:	Document Type Definitions
DTIC:	Defense Technical Information Center
DTIC-E:	Directorate for Information Science and Technology
E-mail:	Electronic Mail
ES:	Expert System
GII:	Global Information Infrastructure
GUI:	Graphical User Interface

HTML:	Hypertext Markup Language
IBM:	International Business Machine
LCDR:	Lieutenant Commander
LT:	Lieutenant
MCCRE:	Marine Corps Combat Readiness Evaluation
MCI:	Marine Corps Institute
MCLS:	Marine Corps Lessons Learned Database
MIT:	Massachusetts Institute of Technology
NASA:	National Aeronautics and Space Administration
NCSA:	National Center for Supercomputing Applications
NII:	National Information Infrastructure
NLCCG:	Navy Laboratory/Center Coordinating Group
NPS:	Naval Postgraduate School
NPSNET:	Naval Postgraduate School Network
NSDL:	Naval Service Digital Library
NSF:	National Science Foundation
OPS:	Operations
PARC:	Palo Alto Research Center
PC:	Personal Computer
RAD:	Rapid Application Development
SGML:	Standard Generalized Markup Language
SIDLP:	Stanford University Digital Library Project
SOP:	Standard Operating Procedure
SQL:	Standard Query Language
STIP:	Scientific and Technical Information Program
TAFIM:	Technical Architecture Framework for Information Management
TULIP:	The University Licensing Program

UC:	University of California
UM:	University of Michigan
UMDL:	University of Michigan Digital Library
USA:	United States Army
USAF:	United States Air Force
USMC:	United States Marine Corps
USN:	United States Navy
USO:	Uniform Services Organization
USS:	United States Ship
VIP:	Very Important Person
WAIS:	Wide Area Information System
WATERS:	Wide Area Technical Reports System
WWW:	World Wide Web
XO:	Executive Officer

I. INTRODUCTION

Grasping the scope of resources and breadth of technologies embraced by the world-wide Digital Library (DL) movement presents a daunting challenge. Even experts, who are professionally immersed in the field, confess bewilderment at the intimidating pace of advance. Many individuals find the evolving vernacular as frustrating to track as the diverse visions of the future offered by scores of DL development efforts. Presently (circa 1996), "Digital Library" has become an umbrella-term that conveniently, if inaccurately, encompasses global efforts to enhance compatibility and establish precision access to remote, electronically stored data resources and processing systems. If that definition strikes the reader as broad and imprecise, it is because at this stage, the Digital Library movement is immature, ill-defined and unpredictable. As this thesis will attest, it is also a field of unlimited opportunity, innovation and excitement.

As we approach the millennium, it is especially appropriate that we find ourselves on the cusp of achieving virtual connectivity to the world's cache of data-stores, knowledge bases, text & imagery bases, Decision Support Systems and Expert Systems, as well as establishing a profound connection between human beings that overcomes the barriers of language, location and politics. Make no mistake, the Digital Library movement is a juggernaut. The stakes for governments, industry, academia and individuals are sufficient to motivate and sustain an immense investment of time, resources and effort. From this collective pursuit of self-interest, the global Digital Library will emerge, configured to meet the needs of those who were involved in its conception and birth. We are convinced that it is in the best interest of the Naval Service to participate at this formative stage, where our contributions can help shape the future.

A. PURPOSE AND OBJECTIVES

The authors present this thesis for the purpose of being used as a comprehensive reference guide for conducting Digital Library (DL) research and initiating the development of a Digital Library to meet the needs of the Naval Service. Our objectives are to stimulate interest in DL technologies within both the Navy and Marine Corps, and define a strategy by which a Naval Service Digital Library (NSDL) can become reality.

B. STRUCTURE AND APPROACH

This thesis contains useful information for individuals interested in the Digital Library movement. As a source of orientation, it should be particularly valuable to librarians, researchers, NSDL stakeholders and potential funding sponsors. To be informative to such a wide variety of users, it is structured into several modules, which can be referenced whole or in part (suggested users in parenthesis):

- Chapter II: Scenarios From A Digital Fleet (All readers) - Scenarios from a day in the life of a NSDL-supported fleet.
- Chapter III: Background (Initiates to the DL field) - An overview of fundamental DL issues.
- Chapter IV: Survey of Current Initiatives (Researchers/Librarians) - A snapshot of current (circa 1996) DL projects.
- Chapter V: Technological Innovations (Technically-Oriented Users) - A review of key DL technological innovations and claims.
- Chapter VI: Creating The NSDL (NSDL Stakeholders) - Our vision and specific recommendations for developing a NSDL.
- Chapter VII: Conclusion (NSDL Stakeholders) - Further research topics, action items and conclusions.

- Appendix A: Thesis Summary (NSDL Executive Stakeholders) - A thumbnail sketch of the issues and our recommendations for developing a NSDL.

C. FOUNDATION OF WORK

This thesis has been constructed upon a foundation of work conducted by the Naval Postgraduate School (NPS) Digital Library Initiative Team, which was chartered by the NPS Provost to conduct an analysis of the dynamic field of work encompassed by the term, "Digital Library" (DL). During an eight-month period, the team focused upon projecting the DL needs of both the Naval Postgraduate School and the Naval Service. Led by the Chairman of the NPS Computer Science Department, the team's configuration included the Director of the NPS Dudley Knox Library and members of her staff, as well as academicians and research specialists from the NPS Departments of Mathematics, Systems Management and Computer Science as well as active duty (USN & USMC) graduate students (Appendix B).

II. SCENARIOS FROM A DIGITAL FLEET

Scenarios provide an efficient method of encapsulating the important characteristics of a system and presenting its capabilities with a unique and interesting perspective. The following vignettes are offered to project the impact of a Naval Service Digital Library on the lives of service members, both professionally and personally. The reader is encouraged to add his or her own expertise and imagination in mapping the capabilities of an NSDL to meet specific needs.

A. OVERVIEW

If the NSDL is created, its primary purpose will be to serve the needs of active duty Navy and Marine personnel in their search for information that will solve problems, increase knowledge and improve performance. Through the NSDL, service members will have unprecedented access to military, academic, research and public domain sources of information. As a benefit, the NSDL can function as a repository of lessons learned and a clearing house for late-breaking topics. The following scenarios represent a small portion of the capacity such a system needs to influence and enrich our lives.

1. Persian Gulf, Aboard a Cruiser

CDR Mike Tryon, the ship's Supply Officer, has been notified by his XO that on their second day in port, they will be host to a delegation of local VIP's of the Muslim faith. He has been given the responsibility of coordinating a shipboard reception for the guests and their consulate escorts. After contacting the NSDL Help desk, CDR Tryon and the Research Librarian agree to divide their labor and he is given pointers to several cultural resource locations. He directs his head cook to research menu options on-line and then conducts two hours of intense research into

the area's history, culture and traditions, tapping resources from the Library of Congress to the host country's Ministry of Cultural Affairs Web site.

While putting together a briefer for the CO, he downloads a report from the NSDL which contains biographical information on the ruling party, a list of national and religious holidays, a collection of media reports, and a list of national heros and celebrities. CDR Tryon notes that this week is the birthday of one the Royal family's young heirs and recommends that the CO invite the young lad aboard for a birthday celebration. The consulate loves the idea. The cook puts together a blend of traditional and American food and the guests are dazzled and grateful for such hospitality. Sailors are given unprecedented access to the port city, lifelong friendships are forged and the State Department receives a glowing report from the consulate. The next morning, CDR Tryon uploads a copy of his research into the NSDL *regional lessons learned database* and secures for liberty.

2. Mediterranean Sea, Aboard a Nuclear Aircraft Carrier

Chief Petty Officer Susan Smith supervises an aircraft engine repair division on a forward-deployed CVN. A key replacement engine for an F/A-18 has failed its high-power checks. Confronted with an engine shortfall and time-critical tasking, Chief Smith has been challenged to fix the problem. Rather than placing the engine out-of-commission per Standard Operating Procedure (SOP), Chief Smith notifies the ashore maintenance facility of her dilemma and, after receiving waiver authority, downloads the refurbishment procedures. With the assistance of an NSDL Research Librarian, she locates digital schematics from the cognizant authority. In response to the Librarian's query, the manufacturer electronically transmits a video excerpt from an in-house training course. Fourteen hours later the refurbishment is complete and the Chief electronically mails a full report to the Repair Officer and the repair engineers ashore. She also forwards a copy to the NSDL lessons learned archive.

3. Sea of Japan, Aboard a Destroyer

Encouraged by his performance on the advancement-in-rate exam and with the support of his Division Officer, Petty Officer John Jones considers tackling a distance learning course. His dream is to pursue the "Seaman to Admiral" program, but he lacks confidence in his ability to complete college-level work. Using the **NSDL** on-line directory, he reviews several math courses and after evaluating the outline, prerequisites and student comments for an introductory calculus course, decides to enroll. Subsequent to registering for this self-paced program, he downloads the first of nine modules, including text, study guide and practice tests. There is also a graphical software application and several lecture videos available. The **NSDL** Research Librarian puts Petty Officer Jones in touch with an NPS Professor who hosts a math support electronic forum. The Ship's Educational Support Officer posts an announcement on his homepage and four of Petty Officer Jones' shipmates sign-up to make a five-person study nucleus. By CO policy, this permits them to schedule the ship's study hall and also qualifies the group to reserve a dedicated on-line connection.

4. Persian Gulf, Aboard an Aegis Cruiser

During a wardroom discussion, Captain Frank Franklin challenges his junior officers to enter the Naval Institute's WarFighting Essay contest. Inspired, the group contacts the NSDL and downloads the contest rules and the top three articles for each of the last five years. Satisfied that they have a worthy topic, LT Mary Miller posts their outline and a milestone schedule on a newly established group directory. Each member works on the project during their spare time and its progress becomes a hot topic. On-line research is conducted via the NSDL and historical data and imagery is downloaded from the Naval Academy Library. The final multi-media package is

submitted electronically to the Naval Institute, prefaced by a video introduction by Captain Franklin. The judges are impressed.

5. Mediterranean Sea, Aboard an Amphibious Assault Ship

Newly promoted Corporal Ben Banatz has been assigned the key billet of fireteam leader. He has been tasked with ensuring his Marines have completed or are enrolled in required Marine Corps Institute (MCI) courses. He visits the company clerk and they remotely access MCI, view each Marine's record, and download course material and final exams. While online, Corporal Banatz queries the **NSDL**, requesting assistance in locating relevant sandtable exercises and tactical scenarios. With help from a Research Librarian, he downloads recent tactical decision games from the Marine Corps Gazette as well as lesson learned from the MCLS database. That afternoon, his fireteam conducts two sandtable exercises and spends an hour working MCI courses.

During his search, Corporal Banatz discovers a new data archive and downloads recent lessons learned from units undergoing a Combat Readiness Evaluation (MCCRE), which he forwards to his Platoon Sergeant. Late that week, his Platoon receives a 30-minute brief on the subject from First Lieutenant Gearhard, in preparation for next month's upcoming evaluation. The Platoon Commander notes his corporal's performance and schedules a leadership meeting to discuss this innovative approach to training.

6. The Mediterranean Sea, Aboard a Ammunition Resupply Ship

Lance Corporal Billy Billings will be voting in his first national election next month and he is bewildered by the choices. Many of his bunkmates already have strong opinions on the issues and he feels left out during their off-duty discussions. With help from an NSDL Research Librarian, he visits the candidate's homepages and downloads excerpts from several position papers. Intrigued, he continues his search

and finds several non-profit candidate rating sources that provide detailed information on candidate voting records. His shipmates are impressed by Billy's solid input to the next discussion.

7. The Caribbean Sea, Aboard a Fast Frigate

Commander Greg Goodguy is the Executive Officer (XO) of a Fast Frigate on deployment. The ship is making an unscheduled port call. He must decide whether to recommend port liberty for the crew, and if so, whether to encourage families to travel and meet the ship. The XO accesses the Naval Service Digital Library via the Internet and, within minutes, downloads current versions of the CIA fact book and State Department advisories for regional countries, as well as a report filed by an XO whose ship visited the port three months earlier. His request for further information is processed by an NSDL Research Librarian who screens and compiles a list of pointers to relevant sources, including video and image archives at Stanford and CNN. His query triggers a response from the closest USO pointing to their "Welcome Aboard" home page. The Captain of the ship approves the XO's recommendation for liberty and Commander Goodguy posts a complete on-line, multimedia visit guide for the crew and their families, including commercial airline schedules, exchange rates and a list of local hotels. He forwards a synopsis of the port visit to the local consulate through the NSDL E-mail drop and posts a duplicate of his research file in the NSDL regional lessons learned forum. [Ref. 58]

8. Persian Gulf, Aboard a Destroyer

Captain Roy Rogers is conducting wardroom training. His sonar officer is a newly qualified lieutenant who asks, "Sir why doesn't our sensor work properly in this scenario?" The CO agrees that the fleet needs an answer to this long-standing problem. A query via the NSDL reveals three agencies of the Navy Research Lab are working on related projects. He posts a message outlining the problem and possible

solutions to the three labs, the regular chain of command and the Naval Service Digital Library "Hot Topic Clearinghouse."

The problem is noted by the Navy's Design Agent who initiates procedures to modify the specification for the next-generation sensor design. An NPS thesis student notes similarities to her current research and devotes several pages of her thesis to explaining the characteristics of the problem. Meanwhile, the tactics development specialists promulgate a partial work-around to current tactical SOP and add several evaluation scenarios to an upcoming fleet exercise. [Ref. 58]

9. In Port Persian Gulf, Aboard the Same Destroyer

Upon arrival in port, Ensign Edward Everywhere follows the in-port arrival checklist and downloads approved updates to the ship's combat control computers. He reports that the NPSNET virtual environment software includes a new sonar visualization module. Captain Rogers notes this and assigns his sonar officer to evaluate it. The sonar officer visualizes the physical response of his recent sonar question (Scenario 8) as part of a real-time, 3-D multi-player exercise. A playback transcript is critiqued during wardroom training the next day and results are posted in the "From the Fleet" forum at the Naval War College. [Ref. 58]

B. WRAP-UP

The concept of a NSDL is not far-fetched, though there remain many technical challenges. What must be done to pursue this goal, is fairly straightforward. The major libraries and research organizations of the Navy and Marine Corps need a unifying vision that broaches parochial and physical boundaries. A funding sponsor must be identified to provide the seed money for a comprehensive study and development of a prototype system. Finally, there must be an acknowledgment of the need to track and monitor the Digital Library Initiatives external to our own. Chapter VI contains the author's recommendations for both a management strategy and an

organizational framework for developing a NSDL. Chapters III thru V present background information, examine current initiatives and investigate DL technologies.

III. BACKGROUND

The purpose of this section is to acquaint the reader with the concept of a Digital Library (DL) by examining the basis for its demand and the challenges of electronic information management. Following an overview, the chapter is divided into four sections. The first examines the fundamental issues relevant to establishing electronic information resource accessibility; the second contrasts dissimilar approaches toward information acquisition; and the last two introduce the major challenges confronting the advancement of DL technologies and the scope of potential services.

A. OVERVIEW

The movement toward creating publicly accessible, electronically connected, resource-sharing libraries finds its roots in the growing pains of the Internet during the early 1990's. As the population of Internet users and providers rapidly expanded in size, a substantial variance in expectations, capabilities and needs arose. So too, did a collective sense of frustration with the limitations of existing information management technologies.

In the dynamically evolving world of cyberspace, our jargon can be as revealing in metaphor as it is descriptive in content: "Surfing the Net" implies an unstructured, free-wheeling activity that evokes a recreational image. Contrast that with the term, "Information Superhighway," which brings to mind: structure, efficiency, and heavy traffic. The dichotomy between the two highlights the Internet community's dilemma. As a society of information seekers, we want the smoothness and continuity of a highway, but we are confronted with a turbulent and totally unpredictable ride, not yet suited for novices. Accessing and exploiting the Internet

is significantly more difficult than entering a merge lane. For many, the "Banzai Pipeline" conjures a more accurate depiction of the experience.

Despite their diversity and often conflicting goals and objectives, the Internet community seems to have collectively grasped that mere connectivity to electronic resources cannot guarantee utility or satisfaction. Without effective information management, the Information Superhighway will remain an unpaved dream. The search for an appropriate model, upon which to base the enormous task of re-structuring the world's stockpiles of data resources, uncovered the overlooked, and often unappreciated, field of Library Science. Though a promising candidate, at issue was the adaptability of library technologies and practices from the realm of maintaining on-site collections of physical media to the management of remotely stored, electronic resources. While preliminary results from several DL research projects confirmed that the principles of Library Science could be applied to the world of electronic media, they identified a significant void in the capabilities of existing information-related technologies. In 1994, several countries, including the United States, committed their resources to numerous, large-scale, well-funded Digital Library Initiatives. Within a few months, these programs were joined by hundreds, then thousands of local development projects aimed at bringing yesterday's academic, public and private libraries into the 21st Century. Each of these programs has self-motivated goals, but together they contribute to a world-wide Digital Library Movement that is collectively expanding the horizon of technology and science.

B. FUNDAMENTAL ISSUES

Electronic access to an almost unfathomable quantity of data has been facilitated by huge strides in both the technology and availability, at low cost, of communications connectivity. This trend should continue, though not without difficulty. A major obstruction to the attainment of on-line accessibility to remotely

stored data is the requirement for both the user and provider to establish compatibility through standardization. Again, there has been significant progress in the establishment of standards and protocols which have strengthened our ability to tap widely distributed, data-rich resources, as evidenced by the growth of the World-Wide-Web (WWW). Yet, connectivity and compatibility are only two of many challenges that must be overcome before we can efficiently share information across the globe. One major problem is that, though the Internet provides exceptional access to ***data***, users need access to ***information***, a resource that can be surprisingly elusive.

1. Data vs. Information Resources

Recognizing the distinction between data and information is crucial to comprehending the magnitude of the problems being generated by world-wide connectivity. Data consists of facts and figures, stored in bulk, awaiting future use. Data can be considered the raw material of information. As anyone who has ever done their own taxes can attest, the process of sifting through mounds of receipts (data) to isolate and extract an item of use (information) can range in difficulty from tedious to impossible. Too much data can easily overwhelm, even smother, the process it supports.

A short trip on the Information Superhighway via an Internet Web browser demonstrates the point. One of many powerful Internet search engines can use a key word or phrase to sift through thousands of remote sources and deliver to the user a list, of potential candidate items. Without DL technology, the information-seeker is confronted by a data collection whose size, completeness, accuracy and utility is determined by chance. In a test conducted at NPS on 15 Oct 1995, our search using the key word "Pentium," resulted in a list of 947 sources whose composition spanned the gambit from technical material, to media reports, to humorous articles and personal opinion. While sifting through this pile, we found hundreds of duplicate,

dead-end or nonsensical sites that took many hours to eliminate. A lesson learned from using the Internet is that it is relatively simple to accumulate mounds of data, but chasing down valuable information is a non-trivial task.

Clearly, connectivity is a double-edged sword that, while useful in rounding up potential sources, can cut deeply into one's time budget and still provide a less than satisfactory result. This dilemma is encountered on the Internet daily, by millions of information-seekers, and is magnified for fleet users who cannot afford to waste precious time or bandwidth in pursuit of solutions to crucial problems. It is the demand for efficient navigation, selection and retrieval of information, from millions of remote data sources, that has sparked the Digital Library movement [Refs. 7, 13, and 18].

2. Data Structuring

Information is data transformed by format, filtering, analysis and/or accessibility into a product that has value to the user. To facilitate this capability, a would-be information provider must accurately forecast user needs, employ a robust organizational method and be committed to diligent maintenance. One approach, frequently used for large databases, involves the creation of metadata, which is a separate data-set that provides complementary information on the structure, organization, and content of resources, but does not require the cache of the resource itself [Ref. 41]. Similar to a library card catalog, metadata contains a relevant description of the source and material while providing the information-seeker with a convenient environment to search.

Given quality metadata, there still must be an effective process to interface both user and provider (with adequate security), and functionally isolate and extract the desired information from the data store. Then there must be a suitable mechanism to transfer the product without compromising its integrity. With such a system, a pool

of trained users could conceivably tap, search and exploit this one data resource. The reader should gain some appreciation for the magnitude of the challenges facing the DL movement by imagining this effort compounded by millions of potential DL users and data resources, eventually integrated into a "user-friendly," world-wide system.

C. INFORMATION ACQUISITION

The level of effort required to electronically search, locate and capture valuable information is not simply a function of baud rate, as many think. It is determined by the structure of the data collection, the quality of its indexing, the power of the search and retrieval system and the expertise of the user.

Currently Internet searching is metaphorically similar to casting a fishing net. Without knowledge of the form, density and distribution of the objective, the composition and quality of the "catch," is strictly up to chance. In the world of digital data, this means that the info-seeker must manually sort random results, which can range in utility from useful to absurd. The cost in time alone can be enormous and there is no guarantee that an exhaustive search has been accomplished. To solve this problem, the DL community is debating a new electronic information management paradigm which contrasts two dissimilar approaches to capturing information: The **Library Approach**, which replicates the environment and the related processes of a physical library; and the **Unstructured Approach**, which embodies the information search and retrieval techniques used in wide practice on the Internet today [Refs. 10, 11, and 19].

1. Library Approach

Librarians have established a system that consistently satisfies the differing information needs of a widely disparate user group. This has been accomplished by structuring physical media (data) into logically organized and accessible collections and providing extensive cross-referencing through cataloguing and indexing

(metadata). At its essence, a library supports an information search strategy focused upon:

- Evaluating all valid, available sources for candidate items;
- Quickly and automatically eliminating alternatives;
- Acquiring for review only the minimum number of items required to accomplish the task; and
- Providing a feedback channel from user to provider.

Librarians contend that failure to follow such a strategy results in time delays, incomplete research, storage problems, and increased costs. These are precisely the reasons that led the DL community to apply Library Science to the realm of electronic data resources. In the environment of physical media, Librarians have become so effective at their craft, that library customers universally expect to have their information needs met swiftly, effectively and with minimum fuss. Peter Graham, the Electronic Resources Librarian at Rutgers University, in his article "Requirements for the Digital Library," discusses the necessity for applying the structured approach of library science to the inter-networking environment:

> Users' needs will continue to be what they long have been. Users will want information reliably locatable, so that when they go there (whether personally or on the net) they can expect to find what they're looking for. Users will want information easily accessible: the cataloging must be clear and accurate, and the information must be promptly retrievable. Users will expect information to be available that was placed in the library's care a long time ago; and they will expect that the integrity of the information they get from the library will be assured. [Ref. 7]

Unlike a library, where information is "targeted" with great precision, Internet accessibility to electronically stored information currently follows a different strategy.

2. The Unstructured Approach

Contrast the organized and supportive environment of a library with the lack of structure one encounters on the WWW today. Though early cybernauts heralded its freedom from restriction and regulation, the Internet's explosive growth has brought it to the brink of chaos.

When searching for information, most users set an arbitrary limit on the number of items displayed on-screen, which indiscriminately filters most of the candidate sources because of time constraints. It is doubtful that many individuals routinely inspect sites that have been listed beyond the display limit. What remains is a hodge-podge of topics, linked only superficially by the existence of a key word or phrase. The user is left to wade through this jumbled mess as thoroughly as his or her time and patience will allow.

If a likely candidate for electronic transfer (download) is found, the possibility of successful capture and future utility is dependent upon format comparability and user expertise. In most cases the user is "buying a pig in a poke," with little or no guarantee of accuracy or authenticity. Compounding the confusion are millions of user-generated linked-lists which provide pointers to someone's "favorite" sources. In this situation, the reference is likely offered by a well-intentioned, but untrained person who may be providing misleading or erroneous information. Moreover, these personal lists are erratically maintained and rapidly become outdated. Without standards for cataloging and indexing, and given the disparity between user expertise and interests, the WWW landscape has become a maze of conflicting signposts and is replete with duplication, nonsense links and inactive sites.

For users who face connectivity charges, the problem is magnified. Evaluating candidate items on-line is expensive, but downloading useless material creates system management problems and can tie-up important resources. Other problems include:

- There is absolutely no assurance that an exhaustive research on the topic has been accomplished by the user.

- The quality and accuracy of available material varies from excellent to ridiculous.

- Specificity in search criteria is limited by the lack of standards and technology to index and catalog distributed digital material.

To combat these problems, computer and information system specialists and librarians are teaming up to develop full service Digital Libraries which "...accomplish all essential services of traditional libraries and also exploit the well-known advantages of digital storage, searching, and communication [Ref. 6]."

D. DIGITAL LIBRARY CHALLENGES

Information users are demanding real-time access to remote sources of digitized text, still imagery, audio, maps, video and more. Establishing a local repository of this magnitude, to mirror the structure and administration of a conventional library, would not be feasible for most institutions. Replication of just one data source for local control and access is costly and inefficient when compared to on-line search and retrieval. Linking remote sources of properly structured data extends the information horizon of the user. With suitable networking and cooperative administration, various collections throughout the world are being united to comprise a Large-Scale Network of distributed Digital Libraries for resource sharing.

The DL community - including national, corporate, community, and school libraries - face many challenges, primarily technical in nature, but also cultural. One

such problem stems from the way we handle information items in paper form. Levy and Marshall, in their article "Washington's White Horse? A Look at Assumptions Underlying Digital Libraries," view society as a culture which nurtures annotation. For example, most individuals would cease to function effectively without the ability to mark documents. In certain communities, we spend much of our time making remarks and taking notes on paper forms. This usually occurs directly on the document we are viewing. If it is a book or some other publication, we use a "post-it" note or resort to making copies. These rituals increase the value of these paper items to the individual and help model the basis for their personal and shared files. According to the authors, until our culture learns the techniques required to make the same annotations electronically, a complete shift to digital technologies is unlikely in the near future, presenting an interesting challenge to the DL community. [Ref. 10]

The DL movement is not being embraced by some members of the Library profession. Generally speaking, this community values airtight control of resources. There are significant concerns about intellectual property rights and the ability to assure authenticity and accuracy. Frankly, many Librarians see the DL movement as an ominous and unwelcome intrusion. The technology can be intimidating and there is a shared perception that funding for electronic media will inevitably erode support for both the acquisition of physical resources and library construction. Another problem is that the universities with Library Science curricula have been unable to keep pace with emergent DL technologies. Many Librarians are faced with a dilemma in which they work for a Library that is not technologically advanced, but still need training to remain abreast their peers. In our research we have seen evidence that this issue is polarizing the profession. As enthusiastically as some libraries are committing to digitalization, others are resisting with equal vigor. However, even the most strident opponents will admit that reduced budgets and the

demands of their customers make the eventual adaptation of DL technologies inevitable. From our perspective, we believe that those in the Library profession who are leading the movement toward Digital Libraries must acknowledge their responsibility to train and equip their peers. In the headlong rush toward the future, there is a real danger that many talented and experienced Librarians will be left behind.

The technical and administrative challenges of digitizing source collections, adapting cataloguing techniques from physical to electronic media, creating intelligent search and retrieval systems, managing copyright and commerce issues and establishing connectivity and storage standards are daunting. Agencies worldwide have invested tremendous resources in an effort to address these challenges with technological innovations and detailed analysis. A few of the many topics posing challenges include:

- Scalability issues of information resources that are physically distributed.
- Variance in user needs and sophistication.
- Diversity in hardware performance.
- Heterogenous types of information resources created by a wide variety of groups.
- The need for extensibility to add new collections as well as new data types.
- Requirements necessary to avoid data processing overload.
- Authenticity of information sources.
- Security.
- User interface paradigms.

- Copyright issues.
- Persistence of objects in a distributed collection.

Before the DL concept can become the backbone of an information infrastructure, these and many other relevant issues must be addressed. It is widely understood that many of the DL challenges are inter-related. To accomplish the task of creating an integrated electronic data repository with the same convenience of a physical library, problems cannot be solved independently. [Ref s. 1, 12, 22, 26, 29, 30, 31, and 32] Chapter V elaborates on this issue.

E. DIGITAL LIBRARY SERVICES

The term, "Digital Library" has only recently evolved and encompasses much more than digitized electronic media. Other terms that have enjoyed brief, public familiarity are: Virtual, Electronic and Cyber Libraries. Of the four, Electronic Library is probably most accurate in that both analog and digital media will be available along with physical resources, but, for better or worse, the term that stuck was Digital Library [Ref. 19].

In our research, we have found remarkable conceptual variance in the definition of a Digital Library among users, librarians, information managers and scientists. An initiate's first thoughts of a Digital Library usually focus on the digitization of existing text into electronically accessible formats like CD-ROM, but once new technology is blended with imagination, the incredible potential of Digital Libraries push the limits of comprehension.

A key role in any Digital Library will be that of Electronic Research Librarian. This individual will be the resident expert in using the tools of the trade to isolate and capture information. As new DL technologies are implemented and refined and new

resources become accessible, there is a danger of overwhelming the end-user. This research specialist will function as an important human link between user and provider.

At this time, infant Digital Libraries already exceed the capabilities of traditional and newer multi-media repositories by tapping data collected and stored in remote databases, knowledge bases, text bases and the WWW. Tomorrow's DL customers will not only be able to access all forms of archived electronic material, but will engage interactively with the computational models of Decision Support Systems or interrogate Expert Systems to extract tailored, professional advice [Ref. 2]. While it may be impossible, at this stage, to definitively project the limits of future DL contributions to the field of Information Management, the scope and commitment of world-wide DL efforts are vivid testimony of the perceived value of potential benefits.

IV. SURVEY OF CURRENT INITIATIVES

The purpose of this section is to acquaint the reader with the scope of effort currently underway in DL technology research, by highlighting a number of Digital Library projects. It should be emphasized that these projects are literally the tip of the iceberg; there are thousands of smaller DL initiatives, each of them contributing to the growth and evolution of a World Digital Library. The overview introduces the concepts of both a Global and National Information Infrastructure. Following an examination of the National Science Foundation's Digital Library Initiative in the United States, is a discussion of national and DoD DL projects. This chapter is purposely non-technical. Chapter V addresses the challenges and related technical innovations being pursued by this body of research.

A. OVERVIEW

The challenges which inhibit the location and retrieval of relevant, meaningful, and timely information through electronic inter-networking have spurred academic, corporate, and government agencies to join hands in advancing Digital Library technologies. The pursuit of innovative solutions has been boosted by U.S. Government interest, led by Vice President Al Gore's ***National Information Infrastructure*** **(NII)** strategy, and publicized by the national press under the Information Superhighway slogan [Refs. 12, 25, and 39]. Though motivations vary from economic to altruistic, participants share a vision of transparently networking millions of distributed information resources by expanding the application of the fundamental principles and discipline of Library Science to encompass remotely-stored electronic media. With their foresight and commitment, these agencies are strategically positioning emerging Digital Libraries at the center of tomorrow's *Global*

Information Infrastructure (GII), which will seamlessly link users, providers and their resources across a world-wide continuum. [Refs. 12, 25, and 39]

B. LIBRARY SCIENCE INTEGRATION

Research initiatives at universities, government agencies, and corporate research labs have led to significant enhancements in the techniques by which information is transacted electronically. As discussed in Chapter III, these breakthroughs in information exchange, retrieval and collection prescribed a new role for Library Science, as critical issues surfaced in network data standardization, redundancy, cataloging, indexing, preservation and authentication. The scope of this undertaking is enormous. Figure 1 contains many of the topics and issues related to developing a GII based on DL technologies. The list is not all-encompassing, but rather provides a snapshot of the areas of study, contents, features, issues and roles required to make information resources electronically available and remotely accessible [Ref. 19]. As daunting as this list may be, the perceived rewards for overcoming these obstacles are sufficient to motivate a substantial infusion of time, money, effort and resources by a wide spectrum of contributors.

C. NSF/ARPA/NASA DIGITAL LIBRARY INITIATIVE (DLI)

The DLI is a partnership of academia, private industry and government agencies striving to advance the technologies involved in searching, retrieving, and processing over network topologies. This uniquely structured program is being choreographed by the *National Science Foundation* (NSF), through a joint initiative with the *Department of Defense Advanced Research Projects Agency* (ARPA) and the *National Aeronautics and Space Administration* (NASA). Following a nation-wide competition, six projects, centered at universities throughout the United States, became the core of the DLI in 1994. Each program is focused upon developing a

Abstracting	Education-support	Navigation
Accessibility	Electronic publishing	Object-oriented
Agents	Ethnographic study	OCR
Annotation	Filtering	OODB support
Archive	Geographic info systems	Personalization
Billing, charging	Hypermedia	Preservation
Browsing	Hypertext	Privacy
Catalog	Image processing	Publisher library
Classification	Indexing	Repository
Clustering	Information retrieval	Scalability
Commercial service	Intellectual property rights	Searching
Content conversion	Interactive	Security
Copyright clearance	Knowledge base	Sociological study
Courseware	Knowbot	Storage
Database	Library Science	Standard
Diagrams (e.g., CAD)	Mediator	Subscription
Digital video	Multilingual	Sustainability
Discipline-level library	Multimedia stream-playback	Training Support
Distributed processing	Multimedia systems	Usability
Document analysis	Multimodal	Virtual (integration)
Document model	National library	Visualization
Economic study		World-Wide-Web

Figure 1. Areas of Study, Contents, Features, Issues, Roles

functional DL model targeted at a specific, though sometimes overlapping, set of technical challenges [Ref. 51].

The DLI has been identified as a "National Challenge," crucial to the National Information Infrastructure, because of its potential impact on the Nation's economic and technical competitiveness [Refs. 25 and 51]. Should this effort succeed in developing the foundation of a solid information infrastructure, proponents assert that users, of all experience levels, will have the ability to access, manipulate, organize and digest the sum of information placed at their fingertips. A paper released by the office of the Vice President, _National Information Infrastructure Agenda for Action_,

argues that the NII is much more than just a repository of data stores linked by telecommunications. As envisioned, the NII will include, "...a wide and ever-expanding range of equipment including cameras, scanners, keyboards, telephones, fax machines, computers, switches, compact disks, video and audio tape, cable, wire, satellites, optical fiber, transmission lines, microwave nets, switches, televisions, monitors, printers, and much more" [Ref. 39]. The goals of the NII are ambitious and farsighted and depend upon a successful Digital Library Initiative to become reality.

1. Academic/Corporate Partnerships

The DLI seeks to provide meaningful information to a diverse population, with differing needs, by advancing the means by which information is collected, stored and organized. Its strategy is to generate new knowledge, promote innovative thinking, and accelerate the information exchange process as steps toward developing a stable platform for the NII. The university system was selected as the nucleus of the DLI to foster an environment where the discipline of Library Science could be adapted to electronic information management. Proponents of this strategy were convinced that the active participation of experienced academic and research librarians would ensure that the focus remained on defining and meeting user needs. In 1994, $24.4 Million, distributed over four years, was divided among six universities with demonstrated capacity to conduct DL technology research:

- **Stanford University Digital Library Project (SIDLP)** aimed at developing technologies for a single, integrated, and universal library, composed primarily of large, heterogeneous repositories. [Refs. 17 and 29]

- **The University of California, Berkeley CERES System** to provide widespread online public access to environmental information specific to the state of California. [Ref. 30]

- **University of Michigan Digital Library Project (UMDL)** to explore basic issues in the structure and behavior/function of large scale, distributed (but federated), and evolving multimedia information environments. [Refs. 11, 33, and 26]

- **Carnegie Mellon University, The Infomedia Digital Video Library**, developed to provide on-line video access thru intelligent agents and allow for full-content and knowledge-based search and retrieval. [Ref. 28]

- **The University of Illinois, The Interspace Project**, comprised of a digital collection of interlinked documents and databases for use with a NCSA DL-specific WWW browser. Research includes investigation of issues in sociology, semantic retrieval, and the design of future scalable information systems. [Ref. 32]

- **University of California, Santa Barbara Project Alexandria**, developed for providing easy access to maps, images, and pictorial materials relating Santa Barbara, Ventura, and Los Angeles counties with strong research focus in the area of spatially-indexed information. [Refs. 23 and 31]

Selection to receive funding was determined by past performance in DL-related research and the importance and applicability of the target set of problems to the NII. To ensure the success of this highly visible initiative, the NSF chose institutions that were mated with strong industrial partners. This provided additional financial support, manpower, hardware and added stability to the projects. It further ensured that DLI stakeholders included some of the most powerful corporations in the world. A partial list includes: Digital Equipment Corp., Bell Atlantic, Intel Corp., Microsoft, Hewlett-Packard Labs, Xerox PARC, WAIS Inc., AT&T, IBM, Apple Computer and Mcgraw-Hill. [Ref. 51]

2. Parallel Research Initiatives

Less visible research initiatives are also contributing to the advancement of DL technologies. The following efforts are based at universities throughout the United

States and were selected for inclusion in this section because they represent a cross-section of the growing population of DL-related projects.

The *Center for Intelligent Information Retrieval* (CIIR), located at the University of Massachusetts and funded primarily by the NSF, has been investigating DL technologies for three years. The CIIR areas of interest include:

- Advanced retrieval techniques
- Indexing and natural language processing
- Routing and filtering
- Browsing and query formulation
- Text extraction
- Integration with database systems

The lessons learned in these areas of research have been promulgated to other DLI projects. Since the CIIR is backed by partnerships with corporations & government agencies (24 different members), it has functioned as a role model for the university alliance, while continuing to perform a vital role in developing emerging technologies. [Refs. 47 and 52]

To facilitate the mutual exchange of Computer Science Technical Reports, several universities are developing systems to improve current techniques by capitalizing on inter-networking and the WWW. The DIENST project at Cornell University and the WATERS project (Old Dominion University, State University of New York at Buffalo, University of Virginia, Virginia Tech) are but two of many initiatives investigating the development of a means to catalog, index and retrieve reports by using Internet utilities. [Refs. 20 and 21]

The *Corporation for National Research Initiatives* (CNRI) is working with ARPA to integrate the Computer Science Technical Reports projects of five universities as a prototype Digital Library with emphasis on researching key aspects of storage, search, retrieval and display of information. Participants include: Carnegie Mellon, Cornell, Berkeley, Stanford and MIT [Ref. 12].

D. NATIONAL LIBRARIES AND DL TECHNOLOGIES

Besides the NSF DLI, national libraries worldwide have made significant commitments toward integrating inter-networking technology and library science. The U.S. Library of Congress, the British Library, and similar large-scale efforts in Spain, Japan, Australia, Singapore and other countries are pursuing various technical approaches to digitization, storage and retrieval. The realities of rising costs for acquiring and maintaining printed materials, coupled with the demand for enhanced technological capabilities by the user community, have increased economic pressure on already thin budgets. A Mellon foundation study confirms, "the traditional library's mission of creating and maintaining large self-sufficient collections for their users is being threatened." [Ref. 10]

In light of these and other encroaching problems, many national archives have conducted a full-range of studies on resource digitization and most have elected to pursue these technologies. The U.S. Library of Congress raised over $13 million in grants and contributions in 1994 to digitize a portion of their rare books and pictures archive and have successfully converted thousands of printed pages and images into digital format. [Refs. 9 and 19]

The British Library holds over 18 million volumes and is "...one of the world's greatest treasure houses of written information from every age and culture." [Ref. 34] In 1993, it launched a comprehensive program called Initiatives For Access, which linked 20 separate development projects to investigate hardware and software

alternatives for the digitization and networking of a significant portion of their holdings. One highly visible project, that demonstrates the potential value of networking national archives, is the Patent Express Jukebox system. By connecting multiple CD-ROM jukeboxes, it currently provides on-line access to over one million U.S. and U.K. patents with a mean search and retrieval time of under two minutes. [Ref. 34]

Spain's national program is making significant progress in both the digitization and access of rare scientific books and manuscripts, particularly in the field of medicine. In March 1996, The University of Madrid initiated an on-line catalog for many of the resources recently digitized under the *Dioscorides Project*.

Certainly one of the most impressive programs belongs to the Japanese. Their national DL program currently boasts 24 on-line Digital Libraries. This extremely well-organized program links the country's major universities, national research laboratories and government agencies into one cooperative, resource-sharing system. Not surprisingly, Japan hosted the world's first international DL workshop in 1994.

Copyright Laws & Property Rights

As the foundation is being laid for the Global Information Infrastructure, the issues surrounding copyright laws and intellectual property rights cannot be avoided. These thorny problems must be confronted now, in an effort to develop effective solutions. As the walls of the world's physical libraries, which have traditionally provided protection and security for their resources, are electronically breached, the potential for larceny increases significantly [Refs. 5 and 36]. In the United States, the definitive work in this field was published in September 1995 by the Information Infrastructure Task Force (ITTF), chaired by former U.S. Secretary of Commerce Ronald H. Brown. Titled, *Intellectual Property and the National Information Infrastructure*, this work examines the intellectual property implications of the NII

and includes recommendations for changes to the law [Ref. 55]. The Library of Congress serves as the hub for this field of research. To prevent complex copyright issues from slowing the advance of the DLI, publishers, authors, and libraries are jointly addressing property rights concerns. Their common goal is to avoid legal confrontation by anticipating and resolving problems, without disrupting the progress of technological innovation.

E. DOD DL EFFORTS

In the Department of Defense, the United States Air Force (USAF) has taken the lead in advanced digital technology research. The USAF is currently investigating emerging technologies as a means of achieving strategic advantages and enhancing current and future operational capabilities. Their initiative, *SPACECAST* 2020, analyzes the issues of space exploitation to achieve a Global Presence and high-leverage technological capabilities. The USAF has joined the DLI to ensure an overall advantage in the collection, analysis, synthesis, and dissemination of information throughout the Department of Defense (DoD). [Refs. 3 and 4] Appendix C, is a graphical depiction of how the USAF envisions future DoD information flow.

The U.S. Army (USA) has made a massive commitment to digitization as part of its *FORCE XXI* strategic vision: "Digitization will enable the Army of the 21st Century to win the information war and provide deciders, shooters, and supporters the information each needs to make vital decisions...." [Ref. 44] The two-year old *Army Digitization Office* (ADO) is overseeing the execution of a comprehensive master plan for horizontal and vertical integration of Army-wide organizations and resources. To date, they have completed key activities in establishing target architecture, streamlining acquisition, establishing a common operating environment and ensuring compliance with DoD guidance. This year their focus will be on coordinating digitization activities at the Brigade and below organizational level. Though the focus

of Army efforts are in direct support of the battlefield environment, the benefits of non-tactical information are not being ignored. One example is the new repository of electronically stored information at the *Center for Army Lessons Learned* (CALL). This system links users and providers at all Army Posts. CALL provides on-line access to a central data store that houses both corporate knowledge on a wide array of topics and offers links to additional sources of valuable information [Refs. 16 and 44].

Other DoD agencies are pursuing digitization technologies to enhance their data standardization, redundancy, cataloging, indexing, and authentication capabilities. They are investigating several technological innovations being researched throughout the DLI and have funded various efforts on their own. Much of this work is being conducted in support of the DoD's *Scientific and Technical Information Program* (STIP). As a result, the *Defense Technical Information Center* (DTIC), is conducting research into many DL-related areas. One of DTIC's stated missions is to "...enhance end-user access and to find ways to provide the DoD customer with interfaces to accomplish their mission with ease and efficiency." [Ref. 15] As such, DTIC's *Directorate for Information Science and Technology* (DTIC-E) has developed *GOLDEN GATE*, a PC & Macintosh compatible interface that enables inexperienced users to search databases, display results and order documents, along with providing e-mail and Internet access. [Ref. 14]

DL efforts underway in the Naval Service are not yet consolidated under a strategic plan. Certainly there are many local, command-level projects, but these do not constitute a comprehensive, coordinated effort designed to benefit the entire Naval Service. Besides the effort at the Naval Postgraduate School, one organization that is making progress is the *Navy Laboratory/Center Coordinating Group* (NLCCG). This group, which coordinates the activities of the Naval Research Laboratory, has

begun investigating some of the issues relating to the development of Digital Library technologies, but to date, there is no formal tie-in to the DLI. Further, their efforts are targeted to support and modernize STI access for scientists and engineers, not for the broader military population. [Refs. 24 and 27] Of the major military services, the Navy and Marine Corps have made the least progress in the domain of Digital Libraries.

To position itself to attain an overall advantage and exploit current and future Digital Library initiatives, the Naval Service must comprehend the trends and innovations in advanced research efforts, both within and external to the Department of Defense (DoD). With a funding sponsor, a strategic plan and the commitment of key organizations, a Naval Service Digital Library can come to fruition. Chapter VI contains the author's blueprint for capitalizing on this opportunity.

V. DL TECHNOLOGIES

The purpose of this chapter is to acquaint the reader with several areas of technological innovation being explored by researchers involved in the Digital Library Initiative. There are hundreds of research initiatives striving to solve complex problems that must be overcome before Digital Libraries can be seamlessly linked together. We will discuss efforts in the United States that are germane to the Naval Service Digital Library, but have applicability to other large-scale DL development projects. Following the overview are four sections devoted to presenting information in a format that is readily understandable by individuals not yet conversant with DL technologies.

A. OVERVIEW

Grasping the scope and implications of the innovative claims made by the DL research community may be difficult for anyone not acquainted with its issues and technologies. The jargon used to relate technical concepts can be obscure, even intimidating. This barrier limits the exposure of much of this important work to scientists and researchers. The goal of this chapter is to present key technological issues in a form that is both informative and palatable to traditional librarians, likely DL users and potential sponsors.

In Chapter IV we discussed issues common to the DL community. However, each emerging Digital Library faces a unique set of challenges due to differences in structure, user needs, assets, funding and vision. Hence, no two are likely to share identical traits, capabilities and goals. Therefore, research groups are employing different strategies to solve problems. Our task is to present them in a meaningful context. We have chosen to focus on technical issues related to linking DL users to

remote resources and emphasize technologies that impact the development of a Naval Service Digital Library. These four categories are abstractions that facilitate grouping related technological approaches under common themes:

- **Representation and Finding.** The techniques used by the source provider to characterize stored data items and the technologies used to determine the existence and location of a specific repository by the user.

- **Navigation and Retrieval.** The capability to search a data store with specific criteria and the development of systems to capture targeted items for use.

- **User interfaces.** The presentation of information required by the user to converse with the system.

- **Decision Support Technologies.** Information and associated computational procedures that have the ability to support decision making and be supported electronically.

Any Digital Library project faces the dilemma of selecting a suite of technology options to use in designing their system. By monitoring the progress of other large-scale projects, the NSDL can benefit from their experience and avoid repeating mistakes.

B. REPRESENTATION & FINDING

The successful application of Library Science to the Internet requires data source providers to assume responsibility for organizing their resources (***representation***) in a format that is efficient, searchable and compatible. A key service to be provided by Digital Libraries is the ability to locate data sources (***finding***) that contain information of value to their users.

1. Representation

Millions of dollars are being invested annually in efforts to optimize data representation. Data storage is not expensive, but representation, the process by which we identify each particular data item, is time-consuming and can be extremely costly. Effective representation facilitates the ease with which a user can successfully locate, identify and manipulate an item from a large collection. If query performance is erratic or an inordinate amount of time is required to access data and return results, improvements can probably be made in data representation.

As stated in Chapter III, mere access to data cannot, in and of itself, be considered an asset and, in fact, can be detrimental if one's processing capabilities are over-extended. This is very much the case with the World Wide Web (circa 1996). For each data type (text, images, video, audio, numeric, etc.), there are scores of formats which require the user to employ various format-dependent manipulation technologies. An important goal of the Digital Library movement is to establish reliable, easy to use, platform-transparent, machine/user-independent information exchange [Ref. 48]. New data representation technologies and standards are being created to achieve that level of interoperability [Refs. 6 and 12].

a. Representation Standards

The number, variety, size and growth rate of data repositories mirror the changes in the Internet user community. Representation is a fundamental element in database design and structure and not easily upgraded. Most data representation techniques in use today were designed for limited access systems with finite storage capacity. Changing a representation strategy is a major undertaking for the data provider, requiring time, money and, at least, a temporary loss of productivity. Yet the investment in creating enhanced data structure techniques has been extensive, because the perceived benefits outweigh the costs.

Many of the major research initiatives discussed in Chapter IV are exploring data structure and representation within the context of two evolving areas: **Standard Generalized Markup Language (SGML)** and **Metadata/Indexing.** These are data representation techniques, not specific programs or applications, that enhance a user's ability to search and retrieve data. Metadata/Indexing offers a data solution while SGML is document oriented. A source provider who incorporates these techniques enables searches with more discrimination. With increased depth of coverage, specific data items can be targeted within a record or document, greatly enhancing the user's ability to filter extraneous material and pinpoint specific information.

(1) **Metadata.** Metadata is a special data subset that contains a detailed description of the structure and composition of the source data [Ref. 46]. Much like a library card catalog, it makes data independence possible as specific data items can be isolated from among a large group. For this reason, databases employing metadata are called self-describing. Within such a database, metadata is usually stored in systems tables. This approach enables the user to directly query the metadata vice the source data. If a database contains millions of records, but lacks metadata tables, then queries will be inefficient, requiring the user to wade through the entire database to locate desired information. You may have shared this experience when wandering among shelves of rental videos in a store that categorizes any video less than two years old as a new release. With metadata (a video catalog by title and date of release), the user has the ability to query the much smaller metadata tables (this week's releases) and quickly determine availability and location. When employed by a source provider, this technique decreases user search time and enhances system performance. The challenge is establishing standard formats for metadata that satisfy disparate user needs. [Refs. 6, 41, and 46]

(2) **Indexes.** Indexes also improve database performance and accessibility. Unlike metadata, which contains information on source data structure, indexes are related to specific fields within the database. A database developer defines indexes primarily on forecast user requirements. If she knows, for example, that 'author' is a common search criteria, she will probably designate the author field as an index. An index search will interrogate a data subset based on that single field, corresponding to each record, organized by author name. This will increase the performance of the system by allowing a search of the index for pointers (e.g.,all articles by Steve Jones) rather than the entire database, making sorting and data access more efficient.

Indexes are commonly referred to as overhead data. When a data update is processed in the database (e.g., a record is deleted), each index associated with the record (author, title, topic, etc.) must also be updated. This slows the process of data input and modification. [Ref. 46]

(3) **SGML.** The Standard Generalized Markup Language (SGML) is the International Organization for Standardization (ISO) standard for document description [Ref. 48]. It is a powerful, but straightforward type of document structure which enables cross-platform exchange of information items. The SGML is widely accepted throughout the DLI and used in almost every NSF funded project. Since it's based on an ASCII file format, SGML is compatible with virtually all applications and digitized documents. A significant advantage of this approach is the fact that SGML is the foundation for HTML (Hypertext Markup Language), the unofficial standard document representation language of the WWW. Hypertext establishes an environment where document format is customizable and then automated by the user and enables documents to have embedded links to other sources. With this type of document structure, the user can browse through resources

at her own pace, following a unique path, while choosing what to display and what to skip. The growth rate of the WWW is ample testimony to the value of this technology. A *White Paper* released by the Novell Corporation explains why so many companies are leaning toward full standardization of the language:

> Its founders understood that document format would always present a problem and designed SGML to remove the format from the content and structure of a document. Because SGML preserves document structure, the layout and format can be automated. This means that pieces of information from different sources can be assembled, after which format and layout will be added automatically. [Ref. 48]

SGML provides a map of a document by tagging specific data items for search and retrieval. The tags are known as Document Type Definitions (DTD). The DTD are SGML specific syntax that are similar to the HTML tags you may have encountered on the WWW. A memo structured in SGML might begin like this:

<memo>,
<address>To: Steve Jones</address>,
<sender>From: Dave Jacobson </sender>,
<date>Date: May 5, 1995</date>,
<subject>Re: Thesis Requests</subject>,
</memo>

Notice that SGML provides for easy markup by using the same tag syntax for the start and end DTD, adding a forward slash to designate an ending tag. Although there is standard DTD syntax, the language is flexible enough to allow for user-defined DTD. To enable a more robust search capability, the programmer can tailor the language syntax to a product or document.

SGML is commonplace in the publishing community since it preserves the structure of a document. This ensures that the document remains intact during processing and information content is not accidentally altered or deleted.

b. Metadata and Index Research

The DLI is inundated with research efforts in developing data resources represented by metadata and/or indexes. Rather than storing within database systems tables, many initiatives are placing them on special computers called servers. A server is nothing more than a networked computer programmed to respond to requests from remote computers (clients). Web pages are stored on servers. When a client (Web Browser) requests a Web page using the proper protocol, the server transmits (serves) the requested data. Two examples of this type of technology in the DL realm are the "ComMentor" project at Stanford University and the Wide Area Technical Report System (WATERS) project at Virginia Tech University.

(1) **ComMentor Project.** The basic architecture of the ComMentor project, a research initiative in support of the DLI program, is shown in Figure 2. As shown in the graphic, users interact with the document synthesis

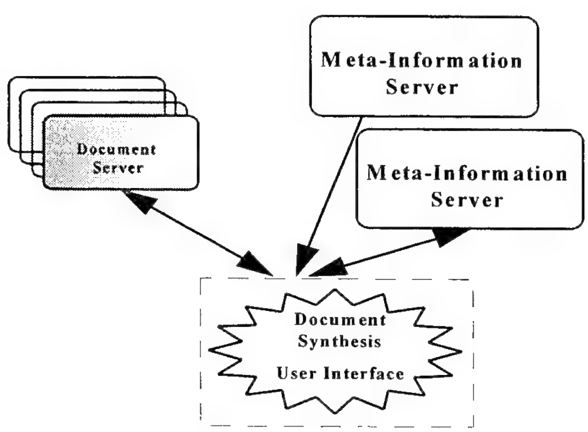


Figure 2. Stanford University ComMentor Project

module, itself on a server, which processes queries to the local meta-information server. Upon request, the meta-information server provides a pointer to information stored on the document server. This type of modularity makes sense in a system designed to handle a lot of traffic. Each server can be configured to maximize performance and the design makes troubleshooting and maintenance easier. If the information requested is not contained in the central archive, the document synthesis module employs Finding Technology to locate an appropriate data source on the WWW. The metadata is then updated with a pointer to the remote site. The cycle continues for each new user. [Ref. 17]

(2) **WATERS Project.** The WATERS project is a computer science technical reports system which uses the Wide Area Information System (WAIS), rather than the WWW, as the basis for its Master Index Server as shown in Figure 3.

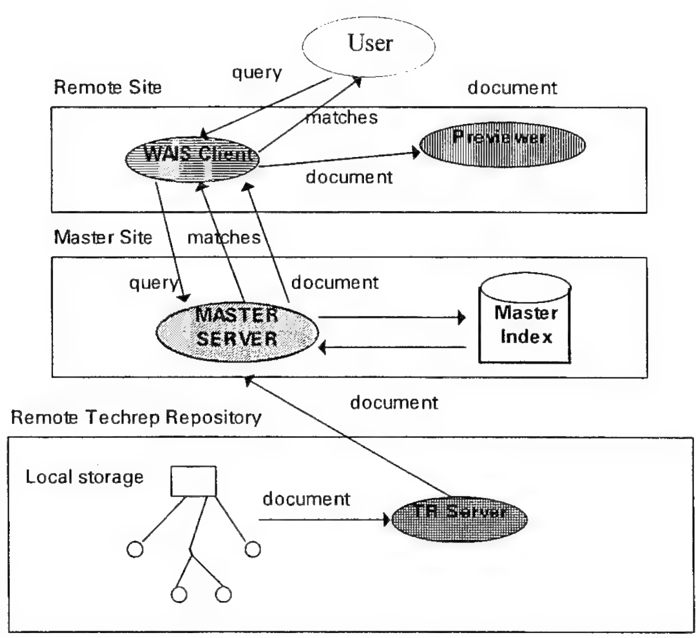


Figure 3. WATERS Project Data Flow Representation

Instead of using metadata, WATERS relies on indexing. When a WAIS client equipped user queries the system, he is searching the WATERS index repository. Pointers are provided to remote databases on the WAIS network. There is no central archive. As long as the Master Index remains up-to-date, WAIS data repositories anywhere in the world can be effectively searched by one-stop-shopping. The user is simply pointed to the data source and establishes a direct connection to retrieve the desired information. The master index will continue to grow in size, but the actual computer science technical reports remain on-site, distributed throughout the world. The only requirement for a provider is to maintain a local index repository that can be interrogated for changes. A small WATERS software program is kept in residence that routinely checks for changes and automatically updates the Master Index. [Ref. 37]

c. SGML Research

The use of SGML has allowed the University of Michigan (UM) to rapidly create a prototype system in support of their DLI program. With the agreement of their major corporate partners, including Mcgraw-Hill and Elsevier, UM has built their DLI archive with forms and full document text in SGML format. This enables their search and retrieval mechanisms to actually penetrate the data allowing users the freedom to customize format. Also based at UM, the **DIRECT (**Desktop Information Resources and Collaboration Technology) and **TULIP** (University Licensing Program) projects are receiving full text documents and images from publishers and contributors in SGML format. [Refs. 11 and 33]

The University of Illinois DLI project has progressed well into their prototype phase by using SGML. We will discuss their project in the next section. Although SGML is a powerful tool, it does have limitations, including:

- **Implementation of SGML is expensive** - It may take a professional programmer to tag the document as desired.

- **Training individuals to use the language may be difficult** - There is a learning curve involved in getting personnel acquainted with language syntax.

- **Implementation takes time** - The same phases of requirements and design needed in coding software are essential when programming in SGML. This takes time and requires experience.

2. Finding

So far we have discussed how data can be represented to facilitate searching. This is a source provider issue, attacked individually, dependent on the type of service a provider is willing to offer. For users, the problem lies in locating a resource that meets their needs. Most DLI projects use the term *finding* to define this area of research. To query a single document or database is relatively straightforward, but isolating the best data source, from among millions, is extremely difficult. The problem is challenging for these reasons:

- **Scope** - There are countless numbers of repositories worldwide which contain vast quantities of data.

- **Variety** - Repositories contain resources of different types in unique formats which currently require different searching mechanisms.

- **Query Language** - Standard Query Language (SQL) and its derivatives are not yet powerful enough to accurately translate and represent the user's desires, thus limiting the specificity of the search.

Ideally, an exhaustive search of all available resources needs to be done quickly and efficiently. Once the user interface agent, which will be discussed in greater detail later, intelligently translates the user's needs, there needs to be a mechanism for locating potential information resources. The SIDLP defines this

effort as a Network Finding Service [Ref. 29]. Unlike current search engines on the WWW and WAIS, these finding services must be intelligent enough to locate potential items of interest and conduct a quality check. The return of hundreds of possible candidates, as discussed in Chapter III, does not meet DL finding resolution requirements.

The major DL research projects have attacked this problem from slightly different angles. Stanford University has done an exceptional job of elaborating their vision for a Finding Service [Ref. 29]. They divided it into three main parts:

- **Network Finding Service** - The same general concept, outlined above, as in the ComMentor project. The search begins with the meta-information server and a pointer is returned which indicates the location of the resource.

- **Search for Candidate Sources** - The query is deciphered to keywords (Boolean value) and potential candidate sources are weighed based on historical data contained in meta-information server. The historical data is used to create a histogram of all potential candidates and the user is given a pointer to the most likely one. Not a high-resolution, precision technique, but more robust than search engines in use today.

- **Distributed Servers** - One course of action the SIDLP is investigating involves distributing its Finding Service. To enhance performance and accessibility, the project is dividing the world into regions and making select servers responsible for collecting and maintaining data on their assigned area. Scalability, redundancy, and security are major research areas still under investigation.

Their approach is founded on research conducted by XEROX PARC. To date, Stanford has been quite successful. By 1998, it is hoped that their popular search engine, YAHOO, will incorporate these Finding Service technologies, thus bringing this powerful Digital Library technology to the World Wide Web.

C. NAVIGATION & RETRIEVAL

The rate of progress toward developing a World DL is very much dependent on our ability to create powerful retrieval engines, sometimes called collection-interface agents [Ref. 33]. With these applications, users will be able to sift through millions of data resources to isolate promising candidate sources with a high degree of accuracy. The keyword search method, limited to title, header, and abstract, of today's WWW search engines are not robust enough to be used with tomorrow's DL. Yet, the concept of full-content searching is unrealistic until effective standardization is enforced amongst the varying data archives. In a report generated at a NSF Workshop in late 1994, *Research Priorities for the World-Wide Web,* it was determined that additional incentives must be provided to foster research in the area of information retrieval, especially if the NSF wanted to see Digital Libraries become a reality. The group recommended "...ongoing support for research into 'information retrieval' and the field of hypertext, multimedia systems, and human-computer interaction, especially as they relate to the problem of finding information in large collections...particularly the Digital Library [Ref. 45]."

The Center for Intelligent Information Retrieval (CIIR) and the DLI are supporting several research initiatives that are examining enhanced retrieval techniques and strategies. The majority of projects funded as part of the DLI are still in the design phase and haven't initiated formal testing of their prototype systems, but there are a few programs that have developed prototype information retrieval engines. These retrieval systems are publicly accessible on the WWW, but searching is limited to local archives that are compatibly represented. [Refs. 47 and 52]

To give a flavor of the technologies being developed, we will discuss four research initiatives representative of the work being conducted. These groups are at the forefront of information navigation and retrieval research.

1. Content-Based Full Text Navigation & Retrieval

Content-based full text navigation and retrieval enables a keyword search of an entire document, vice just the title and/or abstract. If this technology becomes feasible for distributed archives, then DL users will have the ability to comprehensively search entire collections for specific information rather than simply identifying likely sources. Standardization of data resource representation (i.e., SGML) presents the greatest challenge. Despite this obstacle, the University of Illinois elected to use this methodology for its DLI project. Supported by a key industrial partner, they have used commercial software products to rapidly move through prototype to a functional full-text navigation and retrieval system.

Dataware Technologies, a commercial software development and consulting organization, contributed their expertise to the enterprise. This company is a pioneer in developing comprehensive information retrieval solutions based on their advanced full text management and retrieval software. Features of this robust software include:

- **Relevance Ranking** - Weighing criteria for keyword searching.
- **Cross Database Searching** - Allowing distributed searching among diverse computer platforms.
- **Saved Searches** - Saving state of query for re-use.
- **Logical Searching** - Enabling searches through word association (i.e., Thesauri).
- **Multimedia and Image Support** - Providing index schemes for multimedia and images which enables the rapid search of related items [Ref. 32].

Not surprisingly, the University of Illinois has advanced far ahead of their DLI peers in this area by electing to go with an off-the-shelf information navigation and retrieval solution. While the other DLI research testbeds are developing software,

the University of Illinois has been populating its database and enhancing its user interface agent.

2. Image Browsing

Due to its complexity, research in this area is in its early stages. Traditionally, image archives rely on manually generated text captions. This technique has serious limitations. For instance, a photograph of The Louvre might be the target of search for a variety of users. One user might be looking for examples of classic French architecture, another for museums and another for famous landmarks. Yet all of them want the same image. Creating a caption that would reflect all of these aspects is extremely difficult and time consuming. Two research programs are investigating techniques for analyzing and capturing the structure of images using Artificial Intelligence.

a. Alexandria Digital Library Image Research

The Alexandria Digital Library (ADL) project has been working on advanced techniques for navigation and retrieval of image archives. The ADL concept entails pre-processing images before archiving and extracting texture information from each image. Specific examples of textured images include: water, agricultural fields, brick walls, etc. The texture information from these photographs are extracted using Gabor filters, which are generated statistically (modulated Gaussian distribution). These filters trace the outlines of the images and store the results with specific indexes in databases. The indexes are then associated with keywords. Searches using specific keywords can search and locate images that contain no header or caption information. [Refs. 23 and 31]

b. NPS Image Research

Research is ongoing at NPS, under the guidance of Dr. Neil Rowe, that applies artificial intelligence to the image search problem. By incorporating

sophisticated image analysis with a caption, the image can be categorized with high resolution. The goal is a system where a user, observing one image, gains access to related images by using an input device to mark interesting features. This type of navigation would be much like using hypertext on the Web. [Ref. 53]

3. Information Filtering

Stanford University has been working on advanced information navigation and retrieval by developing enhanced techniques for query classifications in a distributed server environment. It should be noted that their efforts are focused on locating potential sources of relevant information, as opposed to identifying the individual documents of interest. This project blurs the distinction between Finding and Navigation. The technique is similar to the histogram discussed earlier in the chapter, in which varying libraries (repositories) are measured on word occurrences by transmitting a boolean query to each archive and measuring the number of distinct occurrences. This information is stored in an index server which can be queried for possible source candidates [Ref. 17]. The SIDLP calls the search technique GIOSS and the results so far have been promising.

4. Bayesian Networks

The Bayesian Network concept is part of the INQUIRY project at the CIIR. It employs advanced text representation techniques as the framework for its approach. Dr. W. Bruce Croft, a pioneer in Information Navigation and Retrieval, formulated the concept based on the work of the ARPA TIPSTER project (INQUIRY is a small sub-set of the project). INQUIRY views the process of capturing client requests and retrieving optimal source candidates as a "probabilistic inference process." It compares text representations based on different forms of linguistic and statistical evidence from natural language queries and user interaction. The project includes four processes: document indexing, query processing, query evaluation and relevance

feedback. For the query process to be successful, document representation must conform to the standards set by the retrieval engine. That is why a document indexing process is part of the INQUIRY project. Interestingly, the project retrieval engine includes a feedback mechanism which allows the client to refine the query as many times as necessary. The project has been very successful and is available for use on the WWW. [Refs. 47 and 52]

D. USER INTERFACE

The key to a new system's popularity lies in the user interface. Thousands of lines of code may create a highly functional and efficient application, but if it lacks an intuitive, user-friendly interface, chances are the system will not capture user interest. A Digital Library interface must somehow translate user needs into machine readable code by promoting a non-threatening conversation with the system. Using the telephone system of the United States as a model, we can conclude that DL users will demand an easy-to-use system that is highly reliable, inexpensive and masks its technical complexities. Moreover, users will expect constant product improvement and universal compatibility. Though this topic entails hardware issues such as input devices and displays, we will focus on the software aspects which include the Graphical User Interface (GUI) and the linkages (scripts) between the GUI and the back-end applications (engines) we've already discussed.

The task of integrating a powerful interface poses a challenge for the DL community. Early on, most DLI contributors recognized their inability to compete in program development with commercial software developers. Why program platform-independent software or construct a new WWW browser when commercial vendors are in cut-throat competition for market share? We can find no compelling reason to recommend spending time and money coding a unique Graphical User Interface (GUI), when off-the-shelf products are more cost effective and already

command user respect. A Digital Library that employs a commercial GUI still has discretion in customizing the look and feel of its interface and will benefit from instant compatibility and future vendor improvements.

WWW browsers, such as NCSA Mosaic, Netscape, and Microsoft Internet Explorer, now dominate the on-line GUI market. Figure 4 provides an example.

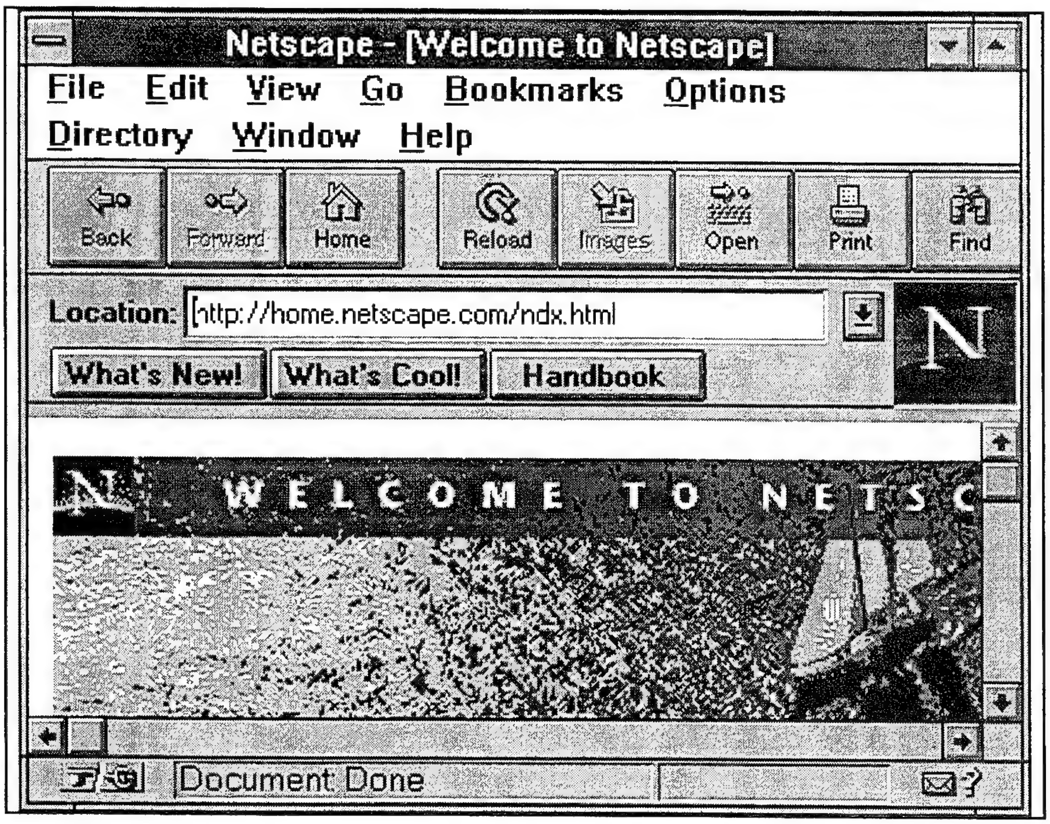


Figure 4. Netscape WWW Browser

While a few DLI projects are stubbornly developing DL-specific GUIs, using Client-Server and Rapid Application Development (RAD) technologies, the majority have shifted to the Internet model which is based on the Hypertext Mark-Up-Language (HTML).

1. HTML

HTML is a simple language that has gained popularity over the past three years; it has been embraced by the Internet community despite not being enforced by rigid standards and not having a governing body. HTML enables anyone to easily publish information on the WWW. The Web pages you've observed on the Internet are written in HTML, yet it limits design and layout options. Much of the sophisticated subject matter you encounter on the Web is actually created in some other, more powerful, development environment. The product is then saved as an image or graphic and imported to the Web page.

HTML is a structuring tool with very basic format capabilities. Its power is in its acceptance as an international de facto standard. HTML's greatest benefit is its simplicity which, ironically, is often cited as its greatest limitation. The quandary facing the DL community and a fundamental challenge to anyone publishing material on the WWW, concerns the level of control the system must exert on its user population. Simply put, "How do you design aesthetically pleasing HTML pages, that effectively extract inputs and return results, and construct an environment where the client wants to stay within system boundaries?"

a. Design Artists

Rather than attack this problem with a technical programming solution, many DLI projects are employing the services of graphical and HTML Design Artists. These individuals have experience creating aesthetically pleasing environments, be they physical or virtual. They are adept at designing HTML pages that engage the

user in meaningful dialog and can present results (output) in a format that is both useful and easily understood. The Stanford Digital Library has constructed a prototype of this type of interface.

b. Interface Example: Stanford DL Interface

Figure 5 displays a user input form written in HTML that captures user input to build a suitable query. Figure 6 displays the results of the search. In this system, books, articles, and other information can be read on-line or downloaded. Copyright laws are enforced by a University Campus "intra-net." Using Z39.50, a U.S. national standard protocol for computer-to-computer data exchange in the TCP/IP network environment, the system maintains connectivity via the Internet. Although the HTML pages depicted are not overwhelmingly exciting, the exchange between user and system is more intuitive than WWW search engines currently in use. Researchers, scientists and programmers have provided an environment where artists can ply their craft to present systems whose appeal matches their functionality. This step is crucial in the evolution of the Internet as it expands the community's resources of imagination and creativity.

2. Linking the GUI to Back-End Applications

The challenge remains to integrate the back-end applications (engines) that provide functionality. These engines are linked to the GUI via **scripts** which are small programs resident on the server, activated by the user through the GUI, containing instructions for data transfer, application execution, display control, etc. Thanks to new products like Delphi, WebHub and Cold Fusion, script programming, at least in the PC world, is fast leaving the domain of professional programmers and being performed by anyone with an interest in adding functionality to their Web site. Appendix D, obtained from DBMS magazine, provides a more detailed list of

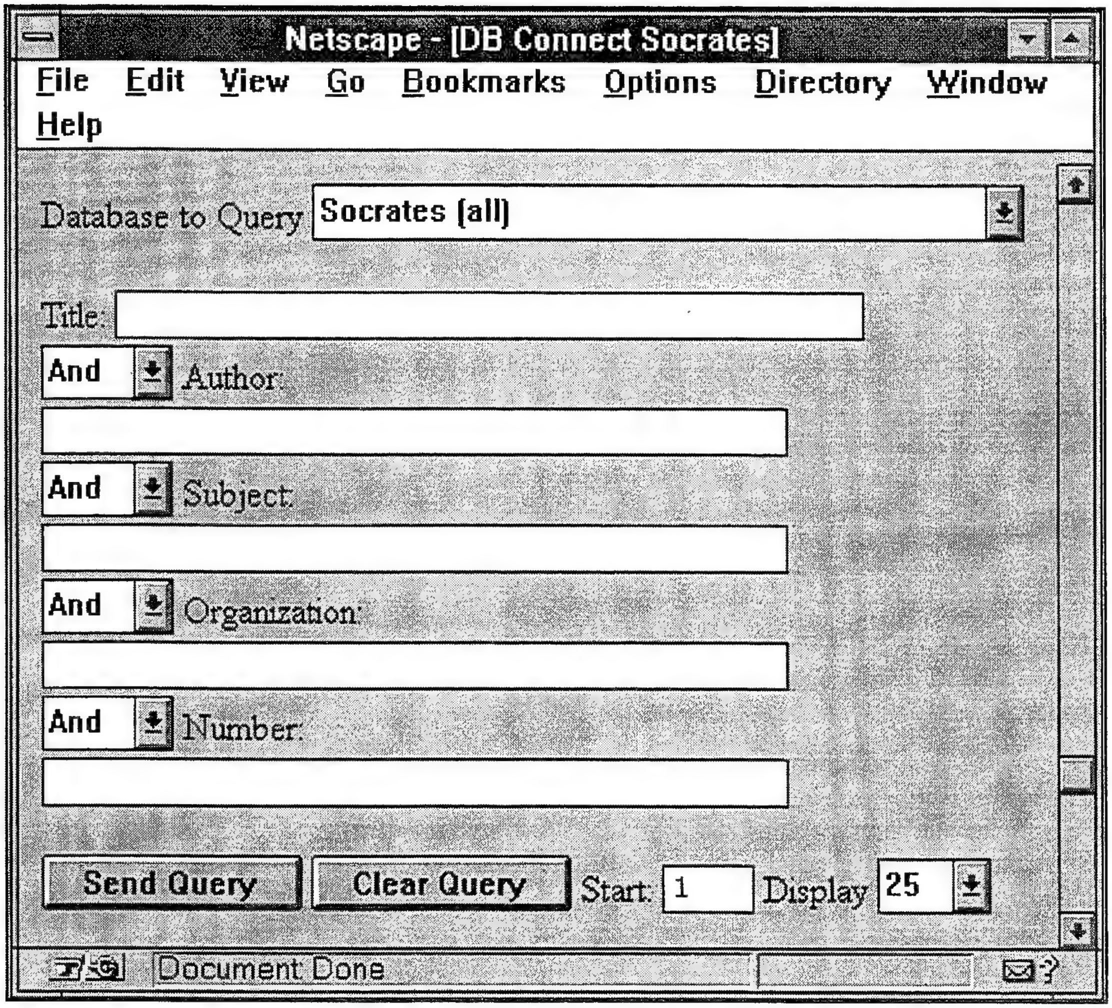


Figure 5. Traditional Digital Library Interface

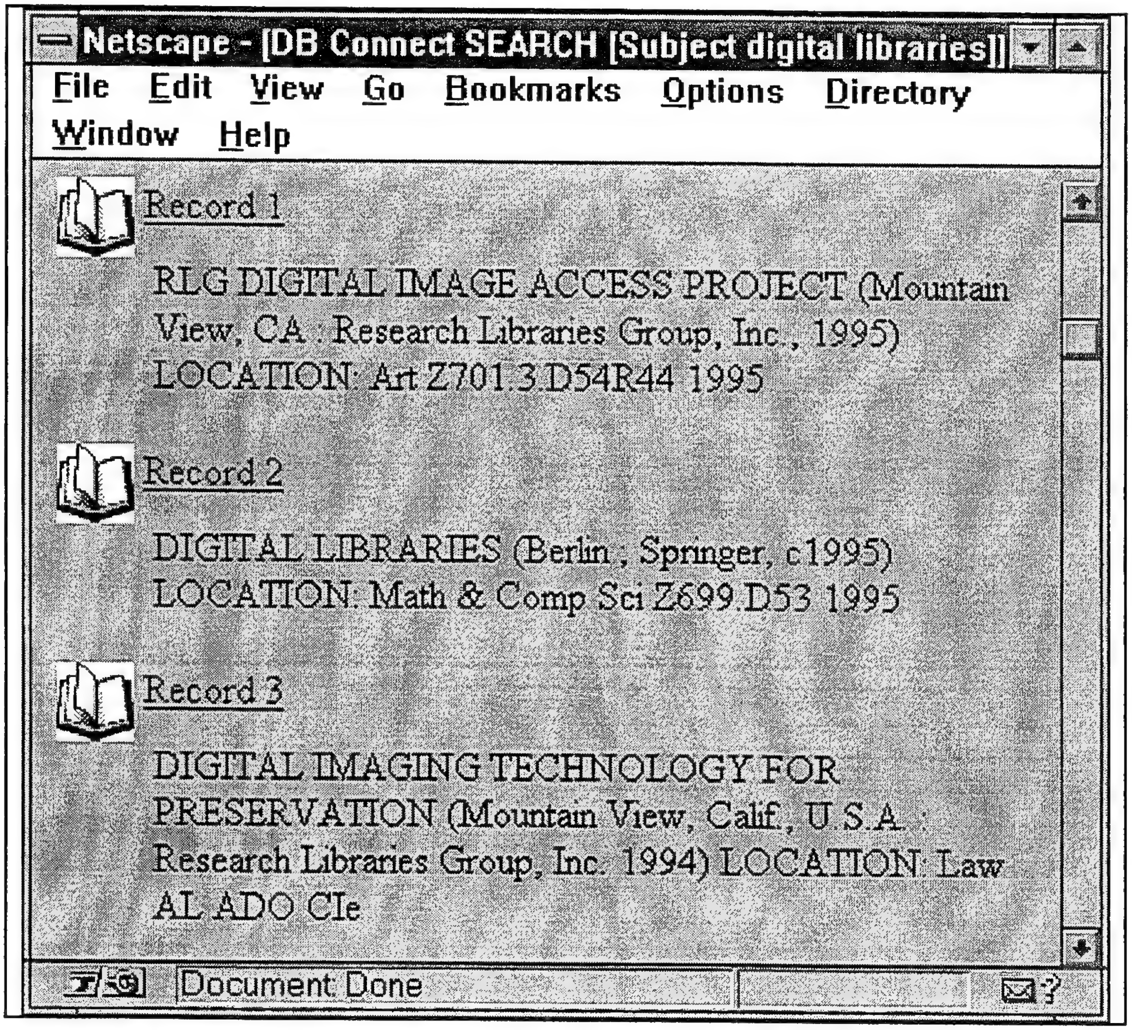


Figure 6. Query Result Interface

resources available for solving WWW-CGI-Database connectivity issue which are confronting the DL community.

Since the preponderance, if not all, of the World DL community will have the capability to access information resource centers via the WWW, the use of Internet utilities to establish user interface makes sense. Until 1995, this technology was based almost exclusively on the premise of WWW clients (users) communicating with HyperText Transfer Protocol (HTTP) servers (providers) that employ Common Gateway Interface (CGI) scripts and/or binary code to access a database or processing system. With this technology, computation is performed remotely on the server with the results being posted back to the client upon completion. As tasks become more complex and demand for services grows, a burden is placed on the provider to design and support a server architecture with sufficient capabilities and capacity. With the release of Sun Microsystems new JAVA programming language, there has been a noticeable shift toward client-side computing, where the server load-sheds some of the work to the client's idle processor. This concept offers tremendous flexibility and may eventually be a cost effective solution to meet many DL user needs. We will discuss both technologies.

a. Common Gateway Interface

The Common Gateway Interface (CGI) is a standard protocol for accessing back-end applications (engines) on HTTP and Web servers. Integrated with an HTML page, which on its own does not interact with the client, the CGI script processes requests and returns a pre-programmed or user-driven output. It is typically an executable program written in PERL, C, C++, Visual Basic, AppleScript, Unix Shell, Delphi, etc., that interacts with a database, text file, and/or computational model. In most cases, the functionality of the executable program and the anticipated user load dictate the server hardware configuration.

CGI requires server-side computing. When a client inputs information into an HTML form and submits a request, a CGI script is typically executed and processed on that server, though tasks can be distributed to other computers. In simple terms, the script transmits data by assigning values to program variables. Depending on what functionality was programmed, a CGI script can perform any number of tasks. The following list, though not comprehensive, provides a glimpse into the power of CGI technology:

- Launch back-end files or programs (i.e., .exe, .tar, etc.) on the same system or distributing tasks among multiple systems.
- Pass values to databases, spreadsheets, or any other program.
- Query databases and files with a Standard Query Language (SQL) and return the output.
- Create animation.
- Send electronic mail.

Researchers developing the DL infrastructure are using CGI scripts extensively. Figure 7 depicts the concept.

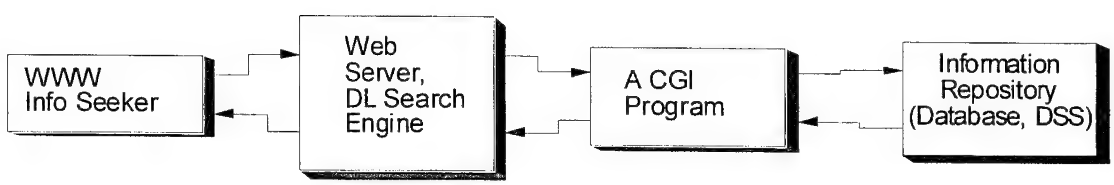


Figure 7. Common Gateway Interface (CGI) Flow Chart

Many of the computation models and intelligent agents that will provide high resolution search capabilities will use CGI scripts as their link to the GUI. The scripts may reside external to the program or may even be hard-coded directly into the agent.

Since server-side computing can require extensive infrastructure to handle client requests, CGI technology has limitations, particularly in large-scale, high-use systems. If the technical challenges can be overcome, client-side computing represents a more scalable solution.

b. JAVA

Sun's JAVA language has emerged as an industry-recognized language for "programming the Internet." Sun defines JAVA as:

> **a simple, object-oriented, distributed, interpreted, robust, secure, architecture-neutral, portable, high-performance, multi threaded, dynamic, buzzword-compliant, general-purpose programming language. JAVA supports programming for the Internet in the form of platform-independent JAVA applets [Ref. 50].**

JAVA applets are small, specialized applications that enable developers to add "interactive content" to Web documents (e.g., simple animation, page adornments, basic games, etc.) [Ref.50]. Applets execute within a JAVA-compatible browser (e.g., Netscape Navigator and HotJAVA) by copying a small chunk of code from the server to client. The applet is processed on the client machine which takes the processing burden off of the server. The language enables both client and server side computing. Figure 8 is a graphical depiction.

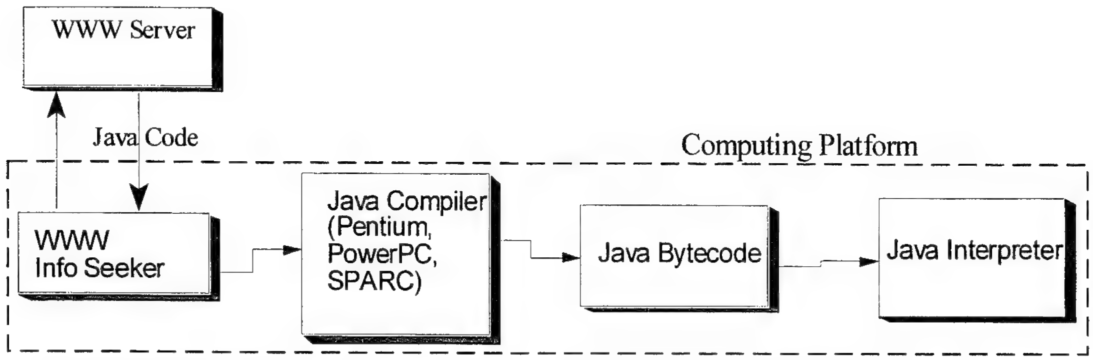


Figure 8. JAVA and Client-Side Computing

Since the JAVA language is in its infancy, it has not yet been incorporated into any of the intelligent search engines of the major DL contributors. However, its potential for reshaping the Internet is recognized by the major vendors, all of whom have established JAVA compatibility in their browsers. One tremendous benefit of this technology is in solving the version control problem that now plagues the Internet. Client browsers come in all shapes and sizes, it is not unusual to have five or six variants of a popular browser such as Netscape in distribution at one time. Each variant has distinct capabilities and limitations which presents the Web Site manager with a dilemma. Do you program for the most or least capable browser, or something in between? With JAVA, the applet contains the code required to execute the program. A browser must only be JAVA compliant, which doesn't limit it from being unique from its competitors. We expect JAVA technology to precipitate a shift toward client-side computing, where the client downloads a platform independent applet and an intelligent agent guides him through the resources of the NSDL to quickly and precisely target information.

E. DECISION SUPPORT TECHNOLOGIES

Research in DL technologies has not been limited to capturing user query information and returning HTML pages with data as discussed in the previous sections. An exciting development, which will provide real-world decision making solutions to end-users (consumers) through Decision Supprt Systems and Expert Systems, has emerged as a new and powerful DL capability. The ability to remotely access powerful decision technologies enables consumers, located anywhere in the world, to solve any number of problems that can be mathematically modeled. Digital Library users represent a huge potential market which should spur development of new systems as providers anticipate the opportunity to recoup costs. Potential users include: a doctor interacting with medical diagnostic and treatment models; a network engineer simulating network capacity and throughput; a logistician planning support for a major force deployment; a combat engineer conducting structural failure analysis; and a contracting officer addressing an acquisition problem.

Until recently, these scientific problem-solving and model-based methods for facilitating decisions were not remotely accessible, limiting their utility to a select group of users. By using Internet technologies to provide an interactive medium, the availability and productivity of these systems is expanded significantly. Research is ongoing at NPS, under the guidance of **Dr. Hemant Bhargava**, to provide WWW access to Decision Support technology. Termed ***DecisionNet***, this project features all types of decision technologies from data sets to modeling environments. Each is registered and made available to consumers. A registered user can search the DecisionNet yellow pages for an appropriate service. The entries in the yellow pages are hyperlinks to the service provider and translates (encodes) the appropriate access semantics so that user and provider can interact. DecisionNet is tailor-made for the Digital Library customer, providing ready access to sophisticated decision tech-

nologies that are no longer isolated by lack of connectivity and standardization. [Ref. 56]

F. WRAP-UP

In this chapter, we've examined DL technologies within four categories for convenience of understanding: Representation & Finding, Navigation & Retrieval, User Interfaces, and Decision Support Technologies. The mechanisms needed to bridge the gap between user and information are in their infancy. Policy issues governing their implementation and use remain undefined. While some DL proponents recommend widespread standardization of resources, using standards like SGML, others claim that search engines can achieve platform independence using new technologies. While it is too soon to tell, we suspect that a hybrid strategy will emerge that encompasses both concepts.

The DL movement is based on meeting the needs of its users. As more Digital Libraries come on-line and begin linking their archives, they will collectively represent an ever-increasing source of demand for electronically accessible resources and services. Source providers will convert to new representation techniques when it becomes economically beneficial. Commercial vendors will continue their competition to capture the User Interface market. Researchers will strive to develop Decision Support Technologies that appeal to a mass market. Meanwhile, DL projects will focus on developing independent, but compatible systems while they continue to push the frontiers of science.

VI. CREATING THE NSDL

The purpose of this section is to introduce the concept of the Naval Service Digital Library and propose both a management strategy and an organizational framework upon which the NSDL can be constructed. A summary of projected NSDL characteristics is contained in an overview following two examples of the system in action. Justification for making NPS the locus of effort for the NSDL project precedes a risk assessment to the Naval Service of not committing to NSDL development. Finally, a recommended plan of attack for coordinating this complex and ambitious undertaking is offered.

A. OVERVIEW

Though substantial effort and resources are being applied to meeting the tactical information needs of the Naval Service, there are tremendous economies and value to be reaped in the non-tactical information environment. The Naval Service can immediately benefit by exploiting the technologies and lessons learned from the world-wide DL movement. To date, there is no mechanism in place to accomplish this task.

1. NSDL Fleet Support

Chapter II describes the impact a NSDL can have on the daily life of fleet users through a series of personalized scenarios. In case the reader has not had time to read them, the following two vignettes are included here to demonstrate the usefulness of a NSDL in providing innovative solutions to routine, but challenging problems.

a. Example 1: USMC Afloat Training

Newly promoted Corporal Ben Banatz has been assigned the key billet of fireteam leader. He has been tasked with ensuring his Marines have completed or are enrolled in required Marine Corps Institute (MCI) courses. He visits the company

clerk and they remotely access MCI, view each Marine's record, and download course material and final exams. While online, Corporal Banatz queries the **NSDL**, requesting assistance in locating relevant sandtable exercises and tactical scenarios. With help from a Research Librarian, he downloads recent tactical decision games from the Marine Corps Gazette as well as lesson learned from the MCLS database. That afternoon, his fireteam conducts two sandtable exercises and spends an hour working MCI courses. During his search, Corporal Banatz discovers a new data archive and downloads recent lessons learned from units undergoing a Combat Readiness Evaluation (MCCRE), which he forwards to his Platoon Sergeant. That week, his Platoon receives a 30-minute brief on the subject from Lieutenant Gearhard, in preparation for next month's upcoming evaluation. The Platoon Commander notes his corporal's performance and schedules a leadership meeting to discuss this innovative approach to training.

b. Example 2: Distance Learning

Encouraged by his performance on the advancement-in-rate exam and with the support of his Division Officer, Petty Officer John Jones considers tackling a distance learning course. His dream is to pursue the "Seaman to Admiral" program, but he lacks confidence in his ability to complete college-level work. Using the **NSDL** on-line directory, he reviews several math courses and after evaluating the outline, prerequisites and student comments for an introductory calculus course, decides to enroll. Subsequent to registering for this self-paced program, he downloads the first of nine modules, including text, study guide and practice tests. There is also a graphical software application and several lecture videos available. The **NSDL** Research Librarian puts Petty Officer Jones in touch with an NPS Professor who hosts a math support electronic forum. The Ship's Educational Support Officer posts an announcement on his homepage and four of Petty Officer Jones' shipmates

sign-up to make a five-person study nucleus. By CO policy, this permits them to schedule the ship's study hall and also qualifies the group to reserve a dedicated on-line connection.

2. NSDL Characteristics

Constructing these, and the other scenarios presented in Chapter II, was an enlightening experience for the NPSDL Project Team. Unleashing our imaginations allowed us to personalize the impact of electronic connectivity and, in so doing, each of us were able to distill the complexity of this topic into a meaningful context. Upon analysis of these scenarios, there were some key aspects which outline the shape and function of the NSDL:

- The NSDL is not an entity that can be defined by physical location, rather it is a hybrid system of linked, heterogeneous components, which must contend with both local and remote users.

- The purpose of the NSDL is to support user-needs, which span an enormous spectrum. User needs will evolve as more resources become available.

- The NSDL has the potential to touch the daily lives of every service-member, in ways that are advantageous and enlightening, both professionally and personally.

- The NSDL represents a quantum shift, from traditional geo-centered libraries, to electronically-connected, information resource centers that must have built-in flexibility, adaptability and scalability.

B. NSDL CENTER OF GRAVITY

The center of gravity for the development of the NSDL should be positioned where information technology users, researchers and providers coexist in an environment conducive to collaborative effort. The **Naval Postgraduate School** is uniquely positioned to lead this effort by virtue of its ongoing research in DL

technologies, joint-service, multi-national student body, renowned faculty, cutting-edge research programs and the Dudley Knox Library facility and staff. Further, its geographic location facilitates close ties with half of the core Digital Library Initiative projects discussed in Chapter IV, which are conveniently located at Berkeley, Stanford and UC Santa Barbara, CA.

C. RISK ASSESSMENT

Should the Naval Service fail to expeditiously initiate its own DL initiative, the results could be costly. A likely scenario would find the Naval Service, in just a couple years, sacrificing a disproportionate amount of scarce resources in a scramble to catch up to its sister services and the rest of the world in the development of a DL infrastructure. The movement toward a global Digital Library is relentless, even within the Navy and Marine Corps. At this moment, functional commands are committing resources, time and effort to establishing or improving connectivity in their efforts to isolate and capture, or provide, information.

The demand from the fleet, as demonstrated in the vignettes, for Digital Library services will increase as improved DL technologies become available to the public and our sister services. This demand will spur significant but, as yet, uncoordinated effort to meet perceived need. The risks of not establishing a structured approach to the development of a Naval Digital Library include:

- Without guidance from professionally trained and suitably equipped experts, who are appraised of the latest DL trends and technologies, commands will waste substantial time, money and effort in an uphill battle to satisfy their information needs.

- Without effective indexing and cataloging, data sources will be useful to a limited pool of users. This polarization of resources will contribute toward resource fragmentation vice unification.

- By not participating in the global DL movement, the Naval Service will not be positioned to exploit advances in DL technology.
- Without an accessible and useful repository of "Lessons Learned," mistakes are bound to be repeated, effort duplicated and time lost.
- Unless the Naval Service aggressively pursues DL technology, its current library infrastructure will rapidly fall further behind the world DL community and its librarians will be placed in professional jeopardy.

It is the opinion of the authors that, should the Navy choose to remain on the sideline during these next few, critically formative years, the configuration of a DoD Digital Library will likely be shaped to meet the needs of the Air Force. The USAF Air University was formally empowered by the Chief of Staff of the Air Force, in Oct 1994, to "facilitate national collaboration" via electronic connectivity within DoD and with civilian institutions [Refs. 3 and 4]. Their Digital Library efforts are integrated into a long-range plan in support of the Air Force's chartered vision statement for the next 25 years: SPACECAST 2020. For a graphical depiction of the evolving DoD Digital Library from the perspective of the Air Force, see Appendix C.

D. MANAGEMENT STRATEGY

To realize the goal of developing a Naval Service Digital Library, there must be a coordinated effort to accomplish the following tasks:

- Characterize the present system (baseline assessment).
- Conduct research to project future requirements and trends.
- Define desired capabilities and applications (target architecture).
- Develop alternative options (migration paths).
- Establish rules and criteria for choosing the best course of action.

For a project as ambitious and complex as NSDL, there should be a structured management approach that has the capacity to capture all relevant data and, through its methodology, distill a myriad of disparate factors into a rational form. With this capability, management has the capacity to make informed and logical decisions.

The NPS Information Management Technology curriculum includes a graduate course, designed by **Professor Carl R. Jones**, that focuses on the DoD Technical Architecture Framework for Information Management (TAFIM). In 1994, Defense Information Systems Agency (DISA) published the TAFIM as an eight-volume set that provides strategic guidance on developing the technical infrastructure of future DoD information systems. Bold in scope, the TAFIM represents a comprehensive compilation of Information Technology Management principles, procedures and guidelines. The TAFIM process invokes a rigorously thorough and well-defined plan of attack for moving from analysis, through design and into execution. It is particularly well-suited to highly complex and diverse problem environments like the NSDL and is recommended by the authors as the management foundation for this ambitious undertaking. [Refs. 54 and 57]

The TAFIM Approach

The TAFIM focuses upon developing an IS technical architecture that both meets the user's needs and is compatible with all other DoD systems. A comprehensive project development lifecycle is defined where coordinating efforts can be synchronized as each step in the process unfolds. Figure 9 depicts the eight sequential steps of the TAFIM process. While each step is important, the key milestones, emphasized by Prof. Jones, are:

- Conduct a Baseline System Assessment with the objective of defining the current environment with emphasis on defining user needs.
- Define the Target Architecture for the optimized information system.

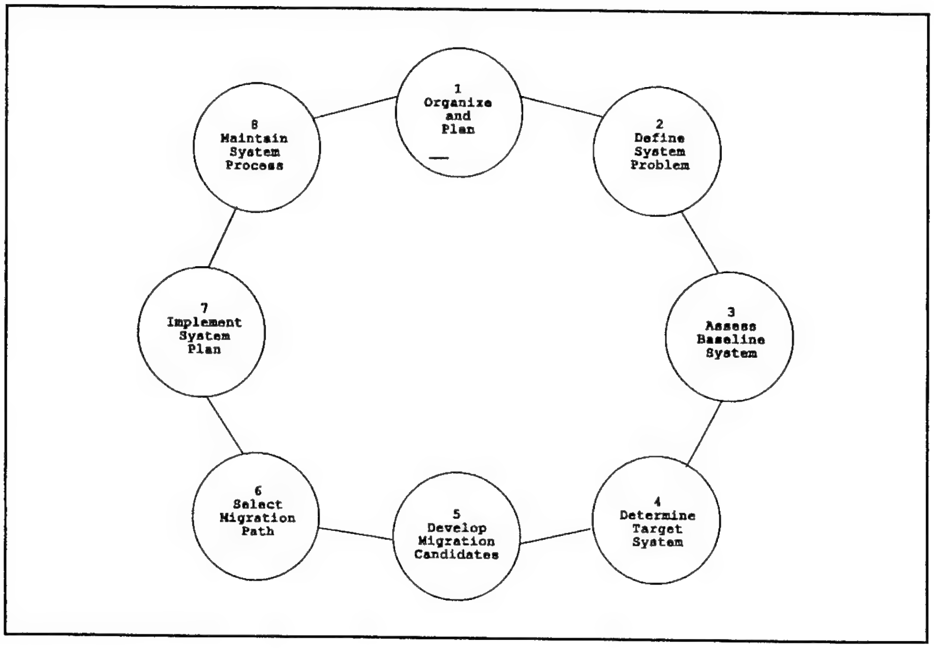


Figure 9. The Structured Approach Process

- Identify alternative Migration Paths which balance conflicting objectives into reasonable tradeoff strategies.

- Establish a fair and suitable method of comparison from which the best migration path can be confidently selected [Ref. 57].

With the TAFIM approach, the NSDL project would have a structure that can support a flexible and dynamic approach. From the outset, coordinated committees could be simultaneously analyzing user needs, tracking other DL projects, targeting and adapting relevant research from within the Naval Service and assessing baseline technical capabilities. The TAFIM is not rigid, but it is comprehensive, and adherence to its recommendations for work flow will provide a significant measure of assurance that vital steps are not omitted.

E. ORGANIZATIONAL STRUCTURE

We propose a plan-of-attack for developing the NSDL that incorporates a seagoing, research DL testbed mated with a collaborative taskforce of key USN and USMC libraries and research organizations. Project milestones would be established by the NSDL Steering Committee. Coordination of effort would be the responsibility of the NSDL Operations Department.

1. Taskforce

The establishment of a NSDL Taskforce will provide a communications framework from which to foster and coordinate collaborative research, define user needs and conduct design efforts. It will take the combined resources and contributions of many organizations to produce a NSDL that effectively fulfills the information needs of the fleet user. Functioning as the lead NSDL organization, NPS could:

- Fence project funding to facilitate creation and administration of a cooperative NSDL taskforce.
- Host NSDL working groups, seminars, conferences.
- Participate in world-wide DL development conferences while representing the needs and perspective of the Naval Service.
- Coordinate collaborative research in DL technologies.
- Provide guidance, expertise and training in DL related matters.

Key NSDL Taskforce participants would include: NPS, Librarian of the Navy, Naval War College, Marine Corps University, Naval Academy and the Naval Research Laboratory.

Once assembled and organized, the first task of the NSDL Taskforce will be to draft a Plan for conducting a comprehensive analysis of the unique DL needs and

requirements of fleet users. The TAFIM provides guidelines for this activity. The NPSDL Project Team suggests an aircraft carrier as a suitable environment for conducting user-needs research.

2. R&D Testbed

User-needs study and design development can best be accomplished through the establishment of a research and development testbed that mirrors real-world constraints and capabilities, and upon which, adaptive DL technologies can be evaluated. For this purpose, it is proposed that the **USS John C. Stennis**, or a suitable replacement, be designated the NSDL testbed. **CVN-74** represents a diverse population of USN and USMC DL technology users living and working in an operational, but not yet deployable environment.

Testbed research activities would focus upon establishing a systematic means of capturing, analyzing and projecting user needs for DL technology. This research would serve as a foundation for system design. However, before the NSDL design efforts can begin, there must be an assessment of both user and provider baseline capabilities. Again, the TAFIM provides guidance for this process.

Given an initial assessment of both capabilities and needs, the project team will next identify a target architecture to represent a system optimized for performance. It is against this "ideal" system that future design tradeoffs and compromises will be evaluated and priorities established. Using the R&D testbed, a prototype system can be developed and fielded. User needs and system performance criteria would be further defined as new technologies are introduced and evaluated in a metered, scientific approach. When performance warrants, other Naval Service DL users will be given access to the system to assess flexibility and scalability issues.

With effective coordination, the growth of the NSDL from concept, through prototype to reality could be closely aligned with the progress of **USS Stennis** toward

combat readiness. Using audio-visual teleconferencing technology, regular dialog between researchers at NPS and key personnel aboard the ship could be inexpensively facilitated. To enhance project continuity, graduate students at NPS could be selected for follow-on billets aboard the **CVN-74** and invited to conduct thesis research as part of the NSDL effort.

3. NSDL Steering Committee and Operations Department

The organizational elements needed to coordinate the NSDL effort must reflect both an academic and operational orientation. To support the efforts of numerous working groups, two NSDL control centers are needed:

a. **NSDL Steering Committee** - Charged with defining policy, goals, establishing milestones, fund-raising and conducting project oversight. This group should be comprised of executive representatives of the major NSDL participating organizations.

b. **NSDL Operations Department** - Responsible for scheduling project activities, budget allocation, tracking progress, configuration management, problem-solving and facilitating communication amongst and between work committees.

Though the concept of an Operations Department is somewhat alien to the academic environment, it is precisely these key functions which will make or break the NSDL project. Mirroring military structure is both prudent and necessary to effectively interface with the testbed and project sponsors. A funding sponsor or the CO of the testbed will not be satisfied with "communications by committee" and the problem-solving skills and operational experience of top-notch military officers will contribute to keeping this project on track. The position of NSDL Operations Officer should be filled with a relatively senior and highly experienced individual, with an

Information Technology Management sub-specialty, who functions in the capacity full-time. Other operations billets can be filled by thesis students.

F. WRAP-UP

By its nature, the Naval Service Digital Library will never become a completed project. Just like a traditional library never stops improving services and modifying its collection, there is a fundamental requirement for the NSDL to continuously adapt to changing user needs and technology. The management strategy and organizational structure defined in this chapter provides a framework from which ongoing and future DL efforts can be collectively organized, effectively coordinated and focused upon the attainment of well constructed goals. Technical challenges confronting the development of a Digital library are discussed in Chapter V.

VII. CONCLUSIONS, FURTHER RESEARCH AND ACTION ITEMS

This thesis introduced the concept of a Digital Library, provided a snapshot (circa 1996) of current DL initiatives and examined key technologies and constraints. Among the fundamental issues, a key concept was that too much data, though being the raw material of information, can overwhelm the process it supports. As information seekers, our ever-increasing access to electronic resources has defined a need for new Information Management practices and technologies. In response, the principles of traditional Library Science are being adapted from the local control of physical media to management of distributed electronic resources.

Globally, thousands of ongoing DL initiatives have been undertaken since 1994. Governments, academic institutions and corporations are contributing to this emerging field of DL research and technology. In the United States, the NSF/ARPA/ NASA DLI represents a consortium of distinct, but related efforts to build prototype DL systems in a spirit of cooperative competition. DL technologies encompass a broad scientific spectrum. DL resources will include static archives of text, imagery, audio, numerical data, etc., and computational models that enable remote users to exploit sophisticated Decision Support Systems and Expert Systems. A key role in any Digital Library will be that of Electronic Research Librarian. This individual will help end-users isolate and capture information while providing an important human link between customer and provider.

Examined in this thesis were technical challenges relevant to developing a Naval Service Digital Library, including data representation, resource location (finding), and data store navigation & retrieval. A blueprint for developing a Naval Service Digital Library was presented along with examples of how such a system can meet the information needs of service members deployed around the world.

A. FURTHER RESEARCH TOPICS AND RECOMMENDED ACTION ITEMS

A broad management strategy and organizational framework for NSDL development efforts are provided in Chapter VI. The following areas require further study and attention:

1. Identify potential funding sponsors/advocates and establish liaison. Preliminary efforts are underway at NPS, but this task needs stronger emphasis. We need expertise in this area.

   Action: The NPS Superintendent's guidance and endorsement are required ASAP.

2. Identify potential members of the NSDL Steering Committee (Chapter VI) and invite them to NPS for preliminary discussion.

   Action: Establish a core group and develop a unifying vision.

3. Staff the NSDL Operations Department. Focus initial efforts on developing a systematic method for monitoring ongoing DL initiatives. Travel funds will be necessary.

   Action: Site visits to Stanford, Berkeley and U.C. Santa Barbara.

   Attendance at Digital Library conferences.

   Subscribe to DL-related Internet-based news groups.

   Start a DL development lessons learned archive.

4. Identify NPS students, faculty and staff to join the NPS NSDL team.

5. Identify topics and sponsors for DL-related NPS student thesis work.

6. Define configuration for initial system platform.

Action: Acquire hardware and software.

Construct a Web Server.

Establish audio-visual teleconferencing capability.

7. Design and establish a NSDL Web Site (see item 4).

8. Establish liaison with the Navy Research Lab to explore areas of common interest and coordinate NSDL-related research efforts. The efforts of the Naval Laboratory/Center Corporate Community are coordinated by The Naval Laboratory/Center Corporate Group (NLCCG) with oversight performed by the Naval Laboratory/Center Oversight Council (NLCOC).

9. Examine the progress of the USAF Air University as a model for some aspects of the NSDL program (see number 3).

10. Seek guidance and support pertaining to DoD-specific DL issues from the Defense Information Systems Agency (DISA) and Defense Technical Information Center (DTIC).

11. Seek potential corporate sponsors/partners for related research. Many corporations, most notably Xerox, participate in several different DL initiatives. They possess a wealth of experience and up-to-date knowledge.

B. CONCLUSIONS

Throughout this thesis we have liberally use terms like, *worldwide, global and national*, to describe the breadth and scope of work being conducted in support of Digital Library development. While technically accurate, these are misleading descriptions if the reader assumes there is some form of sophisticated coordination linking the thousands of separate DL initiatives under a unifying vision and structure. Truthfully, this body of work is so new that its boundaries, vernacular and configuration are changing daily. While there are alliances and coalitions, there is as

much dissension as agreement on such fundamental issues as: vision, goals, purpose, ownership, user rights, property rights and commerce. Meanwhile, the technologies continue to advance.

The closest analogy we can find is the settlement of the Western Territories of the United States in the 1800's when the growth of technology (railroads, telegraph, etc.) outstripped our society's ability to manage our newly accessible resources. While there was certainly a massive demographic movement to the West, it was not an organized advance. More rush than march, this phenomenon consisted of self-motivated individuals and small groups, each seeking to exploit new opportunities. For each gold strike, hundreds of tragic failures occurred. For every new township, there was ecological damage and human suffering. Mistakes happened and were repeated out of ignorance. Yet, despite the costs, the migration continued unabated and many pioneers, who braved the risks, achieved their dreams.

It may be a stretch to compare today's Internet with the Wild West, but there are striking similarities. The original inhabitants of the Internet (scientists and researchers) feel encroached upon and displaced by a hoard of unruly, unappreciative and uninvited newcomers. There is a sense of lawlessness that has replaced the bawdy houses and black-hat villains of yesteryear with graphic pornography, computer viruses and hackers. The environmental damage caused by uncontrolled logging, grazing and mining is mirrored by junk e-mail and the dead-end, duplicate and nonsensical WWW sites that threaten to choke our ability to navigate and locate resources. Burgeoning industrialization is apparent as Internet service providers such as Microsoft, Netscape, CompuServe and America On Line struggle to capture a multi-trillion dollar market. Even a Range-War is imminent, given the Federal Telecommunications Communications deregulation of 1996, as telephone companies and cable service providers cross fence lines in pursuit of customers.

Faced with such a lack of structure, wouldn't it be prudent for the Naval Service to move forward with great caution and avoid the growing pains? A legitimate question, to which we emphatically answer: "No Sir!"

To complete our analogy, we point to the disposition of power, wealth and resources as the Western United States entered the Industrial Revolution. They belonged, almost exclusively, to those farsighted risk-takers who overcame the challenges that befell or intimidated others, by being more prepared, diligent and cunning. Once in place, these entrepreneurs solidified their advantage by charging newcomers to use what, only a short time before, had been free for the taking. It takes little imagination to foresee a similar evolution in our cyber-frontier.

To date, the Navy and Marine Corps have concentrated their efforts on the management and control of tactical information. By its nature, this field of work is extremely security conscious which, in turn, encourages isolation and inhibits flexibility. Invoking the old 80%-20% adage, it is our contention that 80% of our daily information needs, as service members, are non-tactical in nature and unlikely to be well supported by our tightly controlled combat information infrastructure.

The Digital Library movement represents a unique opportunity to meet our non-tactical information needs. Marine Corps and Navy service members need ready access to the world's data repositories and processing systems to conquer tomorrow's challenges. Digital Library users will require expert guidance in how to use them and share their own resources. By committing to the development of our own Digital Library, the Naval Service establishes a conduit through which we can influence policy, exploit new technologies and tap limitless resources. Most importantly, as our service members shape the future, we equip them with powerful tools and the knowledge of how to use them.

APPENDIX A. THESIS SUMMARY

The purpose of this section is to introduce the concept of a Digital Library (DL) and inspire interest for developing a **Naval Service Digital Library** (NSDL) among potential participants and sponsors.

As members of the Naval Postgraduate School (NPS) Digital Library Project Team, the authors participated in a seven-month collaborative effort with NPS faculty and staff to define the future DL needs of the campus. The research, group discussions and innovative concepts developed by members of the project team are the foundation of this paper. Chapters III thru VII of this thesis provide user-oriented information with varying degrees of historical, technological and practical emphasis.

A. OPPORTUNITY WINDOW

The team concluded that the Naval Service can immediately benefit from exploiting the emerging technologies generated by the network of ongoing DL initiatives. Chapter III details these efforts. In justifying the strength of their convictions, the team offered the following assessment:

> **There exists a narrow window of opportunity for the Naval Service to join the vanguard of Universities, Industry and Governments who are collaborating in the monumental task of defining the scope and architecture of the World Digital Library. Active participation at this stage provides the Naval Service with a forum from which to shape the future, as well as a gateway to tremendous benefits.**

B. NAVAL SERVICE DIGITAL LIBRARY (NSDL) SERVICES

Though substantial effort and resources are being applied to meeting the tactical information needs of the Naval Service, there are substantial economies and value to be reaped in the non-tactical information environment. Chapter II depicts

the impact a NSDL can have on the daily life of fleet users through a series of personalized scenarios. The following vignette is included here to convey the usefulness of a NSDL in solving routine, but challenging problems.

1. Scenario: Port Call

Commander Greg Goodguy is the Executive Officer (XO) of a fast frigate on deployment to the Carribean. The ship is making an unscheduled port call. He must decide whether to recommend port liberty for the crew, and if so, whether to encourage families to travel and meet the ship. The XO accesses the Naval Service Digital Library via the Internet and, within minutes, downloads current versions of the CIA fact book and State Department advisories for regional countries, as well as a report filed by an XO whose ship visited the port three months earlier. His request for further information is processed by an NSDL Research Librarian who screens and compiles a list of pointers to relevant sources, including video and image archives at Stanford and CNN. His query triggers a response from the closest USO pointing to their "Welcome Aboard" home page. The Captain of the ship approves the XO's recommendation for liberty and Commander Goodguy posts a complete on-line, multimedia visit guide for the crew and their families, including commercial airline schedules, exchange rates and a list of local hotels. He forwards a synopsis of the port visit to the local consulate through the NSDL E-mail drop and posts a duplicate of his research file in the NSDL regional Lessons Learned forum.

C. PROBLEM ANALYSIS

Electronic access to an almost unfathomable quantity of data has been facilitated by huge strides in the technology and availability of communications connectivity, principally via the Internet. Therein lies a problem. Most people, like CDR Goodguy, need, seek and use information, not data. To grasp the critical role of Digital Libraries, one must understand the difference. As explained more thoroughly

in Chapter III, data is raw material, collected and stored for future use. Information is extracted or derived from data, transformed into something of value to the user. Mere access to data cannot, in and of itself, be considered a boon to productivity. In fact, incoming data can easily overwhelm the process it supports.

A short trip on the Information Superhighway via an Internet Web browser demonstrates the point. Without DL technology, the information-seeker is confronted by a data collection whose size, completeness, accuracy and utility is determined by chance. In a test conducted at NPS on 15 Oct 1995, our search using the key word "Pentium," resulted in a list of 947 sources whose composition spanned the gambit from technical material, to media reports, to humorous articles and personal opinion. A lesson learned by using the Internet is that it is relatively simple to accumulate mounds of data, but chasing down valuable information is a non-trivial task.

This dilemma is encountered daily, by millions of would-be information-seekers and is magnified for fleet users who cannot afford to waste precious time or bandwidth in pursuit of solutions to crucial problems. It is the demand for efficient navigation, selection and retrieval of information, from millions of remote data sources, that has sparked the Digital Library movement.

D. GLOBAL SOLUTION STRATEGY

Librarians are experts in the field of Information Management. They are trained to acquire, catalog, format, index, preserve and otherwise manage information sources. In a library environment, information is targeted with precision, not culled indiscriminately. When we "go to the library," we expect to find what we are looking for, in short order, with a minimum amount of fuss. If the discipline of Library Science, coupled with new technologies, can be applied to the vast data resources now accessible, the availability of information will be extraordinary.

The technical challenges, as discussed in Chapter V, of digitizing source collections, adapting cataloguing techniques from physical to electronic media, creating intelligent search and retrieval systems, managing copyright and commerce issues, all while maintaining compatibility are daunting. However, the scope and commitment of world-wide DL efforts are vivid testimony to the perceived value of potential benefits. Chapter IV provides a survey of the most significant initiatives including the four-year, $24 million NSF/ARPA/NASA collaboration centered at six leading universities, begun in 1994. These linked projects have attracted partners from atop the Fortune 500 including: IBM, Digital, Microsoft, Apple, Xerox and Hewlett Packard. U.S. Government efforts reviewed include the $13 million Library of Congress project as well as the programs of the U.S. Air Force and the U.S. Army.

E. RISK ASSESSMENT

Should the Naval Service fail to expeditiously initiate its own DL initiative, the results could be costly. At this moment, functional commands throughout the Navy and the Marine Corps are diligently working to establish or improve connectivity in their efforts to isolate and capture, or provide, information. As detailed in Chapter VI, the risks of not establishing a structured approach to the development of a Naval Service Digital Library (NSDL) include:

- Without expert guidance on new standards and technologies, commands will waste time, money and effort.

- Without appropriate cataloging and indexing, data sources will be polarized vice pooled.

- Without a repository of corporate knowledge, mistakes will be repeated, effort duplicated and non-optimal decisions made throughout the fleet.

Should the Navy choose to remain on the sideline during these next few, critically formative years, the configuration of a DoD Digital Library will likely be shaped to satisfy the desires of the U. S. Air Force. In October 1994, the USAF Air University was formally empowered by the Chief of Staff of the Air Force, , to "facilitate national collaboration" via electronic connectivity within DoD and with civilian institutions. [Ref.4] For a graphical depiction of the evolving DoD Digital Library from the perspective of the U. S. Air Force, refer to Appendix C.

F. PROPOSED NAVAL SERVICE STRATEGY

Chapter VI defines the author's proposed strategy for developing a NSDL, including the recommendation that the program follow the strategic guidance for developing technical infrastructure contained in the DoD Technical Architecture Framework for Information Management (TAFIM), published in 1994 by the Defense Information Systems Agency. Organizationally, our proposal mates a seagoing research DL testbed with a collaborative taskforce of key USN and USMC libraries and research organizations.

1. NSDL Taskforce

The establishment of a NSDL Taskforce will provide a communications framework from which to foster and coordinate research efforts, define user needs and evaluate system constraints. It will take the combined resources and collaborative contributions of many organizations to produce a NSDL that effectively fulfills the information needs of the fleet user. Functioning as the lead NSDL organization, NPS could:

- Fence project funding to facilitate creation and administration of a cooperative NSDL development effort.
- Host NSDL working groups, seminars, conferences.

- Participate in world-wide DL development conferences.
- Coordinate collaborative research in DL technologies.
- Provide guidance, expertise and training in DL related matters to Naval Service Libraries.

Key NSDL Taskforce participants would likely include: NPS, Librarian of the Navy, Naval War College, Marine Corps University, Naval Academy and the Naval Research Laboratory.

2. Testbed

To determine the architecture and functionality of the NSDL there must be a comprehensive analysis of the unique needs and requirements of fleet users. This can best be accomplished by establishing a developmental testbed that mirrors real-world constraints and capabilities, upon which adaptive DL technologies can be tested and evaluated with user interaction. For this purpose it is proposed that the **USS John C. Stennis**, or a suitable replacement, be designated a NSDL testbed. **CVN 74** represents a diverse population of USN and USMC DL technology users living and working in an operational, but not yet deployable, environment..

G. CONCLUSION

To position itself to exploit current and future DL initiatives while developing a Naval Service Digital Library, the Naval Service must remain abreast current trends in advanced DL research efforts, define the needs of the fleet user and promote collaborative effort. In addition to encapsulating both the evolution and current state of global Digital Library initiatives, this thesis recommends a strategy to achieve each of these objectives.

APPENDIX B. PARTICIPANTS IN NPSDL PROJECT

The following individuals participated in an eight-month (April 1995 - Nov 1995) study of requirements for establishing Digital Library services at the Naval Postgraduate School. During this period, the DL needs of the Naval Service were also examined and various DL development strategies were discussed.

Dr. Ted Lewis, Chairman/Professor NPS Computer Science Department

Dr. Maxine Reneker, Director/Professor NPS Dudley Knox Library

Captain George Zolla USNR, Executive Assistant to NPS Superintendent

Dr. Hemant Bhargava, Associate Professor Systems Management Department

Dr. Neil Rowe, Associate Professor Computer Science Department

Dr. Craig Rasmussen, Associate Professor Department of Mathematics

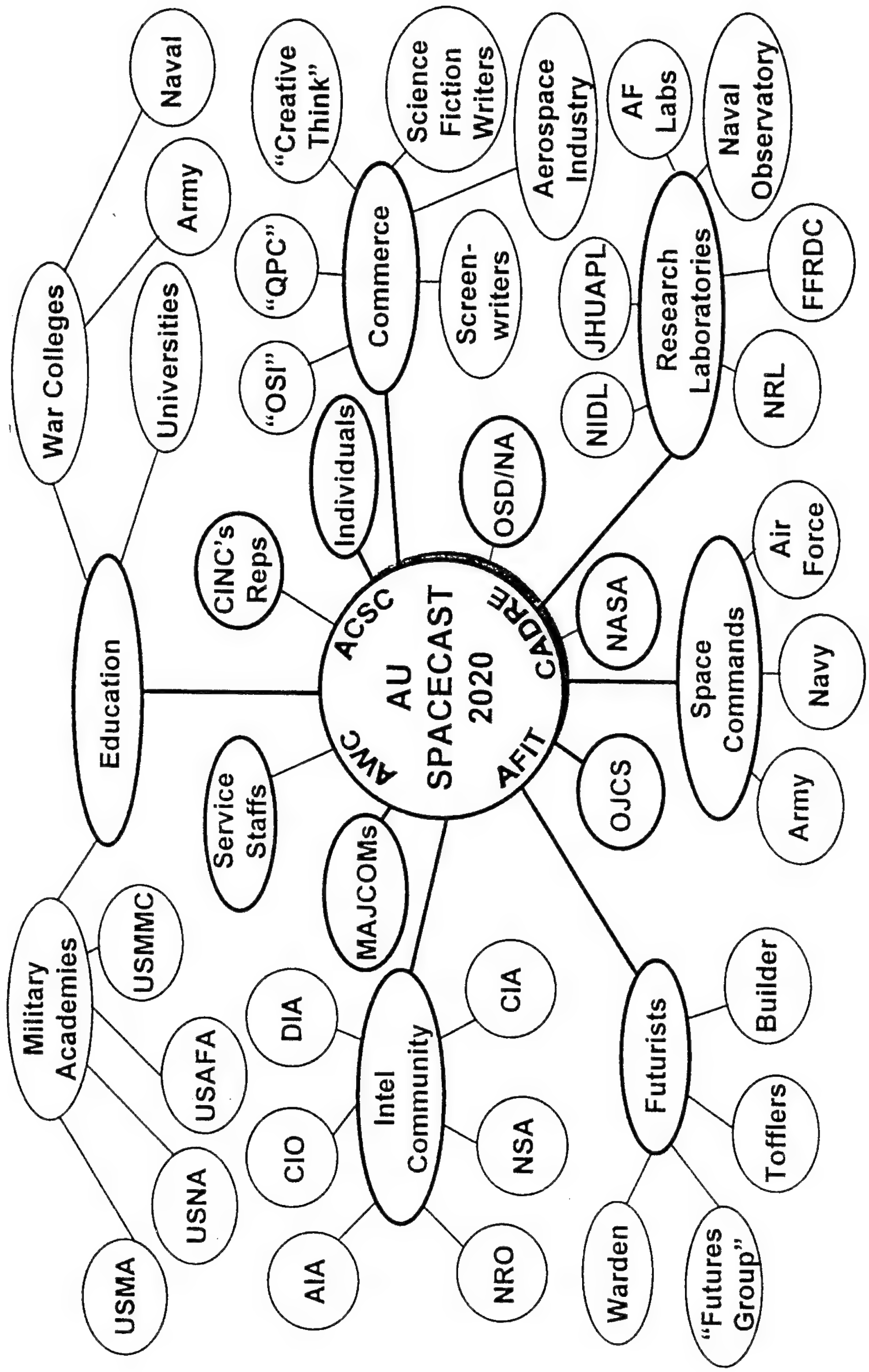
Dr. Don Brutzman, Associate Professor Undersea Warfare Group

Diane Crankshaw, Electronic Resource Librarian, Dudley Knox Library

Commander Bob Norris USNR, Thesis Student (ITM)

Captain David West USMC, Thesis Student (ITM)

APPENDIX C. NETWORK/"WATERING HOLE"



APPENDIX D. WWW CGI TO DATABASE LINK PRODUCT LISTING

PRODUCT/COMPANY & PRODUCT TYPE AND DESCRIPTION

askSam/askSam Systems, Perry, Fla.; 800-800-1997;
904-584-6590; http://www.asksam.com

Version 3.0 of this text DBMS adds support for importing and supporting HTML files into an askSam database.

Cold Fusion/Allaire, L.L.C., Minneapolis,. Minn.;
672-831-1808; http //www.allaire.com

Enables SQL queries by using specialized HTML tags. Uses ODBC.

Concordance/Dataflight Software Inc., Los Angeles, Calif.;
800-421-8398; 310-471-3414; http://www.dataflight.com

A text DBMS with Web publishing features. Concordance can retrieve full-text and fixed-field data stored locally on a LAN or CD-ROM, remotely over the Internet, or at remote locations through a WAN.

Crystal Reports/Crystal Services Inc. (a Seagate
Software co.), Vancouver, BC, Canada; 604-681-3435;
http://www.seagate.com/software/crystal

A report writer enhanced to export reports as HTML pages. and to integrate with CGI programs so reports can be executed from HTML hyperlinks.

DataRamp/Working Set Inc., Lexington. Mass.;
617-863-2339; http://DataRamp.com

Provides access to ODBC data sources over the Web. Includes server and client components in secure (RSA encryption) or "clear" versions.

Dataware Internet Server/Dataware Technologies Inc.,
Cambridge, Mass.; 617-621-0820;
http://www.dataware.com

A gateway that lets Web browsers search Dataware's BRS/Search DBMS (including databases stored on CD-ROM) without any conversion required.

DB2 World Wide Web Connection/IBM Corp., Somers, N.Y.; 800-426-3333; 520-574-4600; http://www.ibm.com

A gateway that integrates Web browsers and HTML forms with IBM's DB2 RDBMS.

---

dbWeb/Aspect Software Engineering Inc., Honolulu; 808-539-3781; http://www.aspectse.com

A gateway using 32-bit ODBC data sources. Also supports Microsoft SQL Server 4.2 and 6.0. Sybase SQL Server 4.2 Oracle version 6 and Oracle7.

---

dbWeb/Axone Services & Development SA, Geneva, Switzerland; +41(22) 342 93 66; http://www.axone.ch/dbWeb

An HTML authoring environment based on Microsoft Access and Visual Basic. It lets Web authors create or generate HTML documents that can be stored in a database or placed on a Web server.

DynaWeb/Electronic Book Technologies Inc., Providence, R.l.; 401-421-9550; http:/lwww.ebt.com

DynaWeb lets Web browsers search and retrieve documents stored in EBT's DynaText electronic book collections. EBT also markets DynaBase, a document repository built on Object Design Inc.'s ObjectStore.

Electronic Workforce/Edify Corp., Santa Clara, Calif.; 800-944-0056; 408-982-2000; http://www.edify.com

Enables Web browsers to use software agents that collect and return data from Btrieve, CA lngres, IBM DB2. Informix. Microsoft SOL Server, Oracle, and Sybase SQL Server databases.

Hype-It 1000, Hype-It 2000, Hype-It 3000 Cykic Software Inc., San Diego, Calif.; 800-295-4295; 619-220-7970; http://www.cvkic.com

A Web server written as an application in Cykic's MultiBase, a relational database end multitasking operating system that supports Xbase programs.

Informix-ESQL/C CGI Interface Kit; Informix-4GL CGI
Interface Kit/Informix Software Inc., Menlo Park, Calif.;
800-331-1763; 415-926-6300; http://www.informix.com

Libraries that let developers write CGI interfaces to Informix databases.

Java/Sun Microsystems Inc., Mountain View, Calif.;
800-821-4643; 415-960-1300; http://www.sun.com

Java lets developers extend the functionality of Web browsers by writing applications a browser can download and execute on the client machine.

KE Texhtml Web Server/Knowledge Engineering
Pty Ltd., Carlton, VIC, Australia; +6-3-9347-8844;
http://www.ke.com.au

A gateway to Knowledge Engineering's KE Texpress ODBMS.

LivePage/The Information Atrium Inc., Waterloo,
Ontario, Canada; 519-885-2181;
http://www.inforium.com/inforium.htm

Stores HTML documents in an Oracle, Sybase, Microsoft SQL Server, or Watcom database. Includes tools to administer, update, and browse LivePage documents stored in SQL databases.

LiveWire and LiveWire Pro/Netscape Communications
Corp., Mountain View, Calif.; 415-528-2555;
http://www.netscape.com

Web application development tools. The Pro version incorporates the Rogue Wave database libraries into the Netscape scripting language, so Web browsers can query Informix, Oracle, Sybase, and Microsoft databases. (A future version will support ODBC data sources.)

02Web/O2 Technology, Palo Alto, Calif.,
415-842-7000; http://www.o2tech.com

A gateway that lets Web browsers access text, multimedia, and complex data from 02's ODBMS.

R:WEB/Microrim, Bellevue, Wash.; 800-628-6990;
206-649-9500; http://www.microrim.com

Add on to R:BASE 5.5 desktop RDBMS. Converts R:BASE forms to HTML and interacts with R:BASE and other ODBC data sources.

Sapphire/Web/Bluestone Inc., Mt. Laurel, N.J.;
609-727-4600; http://www.bluestone.com

Visually creates CGI programs in C or C++ to access Oracle, Sybase, and Informix databases from Web browsers.

Spider/Spider Technologies Inc., Palo Alto, Calif.;
415-969-6665; http://www.w3spider.com

The Spider Development module visually relates HTML form fields to database fields, executes Spider applications, and interacts with an Infonmix, Oracle, or Sybase DBMS and a Web server.

Sybperl/Sybase Inc., Emeryville, Calif.;
800-8-SYBASE; 510-922-3500; http://www.sybase.com

Uses PERL as the CGI scripting language to connect a Web browser to a Sybase database.

Tango/EveryWare Development Corp., Mississauga,
Ontario Canada; 905-819-1173;
http://www.everyware.com

A CGI that integrates Butler SQL with StarNine's WebSTAR Web server, plus a visual editor that creates Web pages that lets you access Butler SQL without writing SQL or HTML code. (An upcoming version will support ODBC access to other data sources.)

Web DataBlade/Illustra Information Technologies Inc.,
Oakland, Calif.; 510-652-8000; http://www.illustra.com

An add-in that integrates Illustra's object-relational DBMS with Web servers. (Illustra's Web site has a searchable version of the DBMS Buyer's Guide.)

WebDBC/Nomad Development Corp., Seattle;
206-448-1956; http://www.ndev.com

CGI gateway using ODBC to pass queries to SQL DBMSs.

WebQuest/Questar Microsystems Inc.,
Woodinville, Wash.; 800-925-2140; 206-487-2627:
http://www.questar.com

A Web server with built-in support for accessing ODBC and SQL data sources.

WebSite/O'Reilly & Associates Inc., Sebatospol, Calif.;
800-998-9938; 707-829-0515; http://website.ora.com

Web server that supports Visual Basic for CGI programs that interact with ODBC data sources and other desktop products.

World Wide Web Interface Kit/Oracle Corp.,
Redwood Shores, Calif.; 415-506-7000;
http://www.oracle.com

A collection of five programs for writing CGI interfaces between Oracle databases and Web browsers.

LIST OF REFERENCES

1. Berners-Lee, T., R. Cailliau, A. Luotonen, H. F. Nielson, and A. Secret, "The World-Wide Web," *Communications of the ACM*, August 1994, Vol. 37, No. 8, p. 82.

2. Bhargava, H. K., R. Krishman, R. Muller. *On Parameterized Transaction Models for Agents in Electronic Markets for Decision* Technologies,WITS '95, July 1995.

3. *The World of 2020 and Alternative Futures, (http://www.au.af.mil/spacecast/).*

4. *Building an Information Infrastructure, (http://www.au.af.mil/ spacecast/).*

5. Doty, P., F. L. Miksa, "Intellectual Realities and the Digital Library," *Proceeding of Digital Libraries '94 Conference,* (College Station, TX, June 19-21), pp. 1-5.

6. Gladney, Henry M., and E. A. Fox, *Digital Library: Gross Structure and Requirements:* from a March 1994 Workshop.

7. Graham, Peter S., *Requirements for a Digital Research Library,* July 1995, *(http://aultnis.rutgers.edu/texts/DRC.html).*

8. webmaster@xerox.com, *XEROX, Inc., (http://www.parc.xerox.com).*

9. Lewis, Peter H., "Library of Congress Offering to Feed Data Super-highway," *New York Times, 12* September 1994.

10. Levy, D. M., and C. C. Marshall, "Going Digital: A look at assumptions underlying digital libraries," *Communications of the ACM, April 1995, Vol. 8, No. 4.*

11. Birmingham, William P., K. M. Drabenstott, C. O. Frost, A. J. Warner, K. Willis, *The University of Michigan Digital Library: This is not your Father's Library,* 1994.

12. Garret, J., "Digital Libraries: The Grand Challenge," *EDUCOM Review,* July 1993, Vol. 28, No. 4.

13. Rand, Roberta Y, "The Global Change Data and Information - Assisted search for Knowledge (GC-ASK) Project," *d-Lib Magazine*, August 1995.

14. *GoldenGate: Graphical User Interface (GUI) for DTIC Systems, (http://www.dtic.dla.mil/index.html).*

15. *Mission of DTIC-E, (http://www.dtic.dla.mil/index.html).*

16. *Army Lessons Learned, (http://call.army.mil:1100/call.html).*

17. Ruscheisen, M., "Content Rating and Other Third Party value added Information Defining an Enabling Platform," *d-Lib Magazine*, August 1995.

18. Lesk, Michael, "The Digital Library: What is it? Why should it be here?" *In Source Book on Digital Libraries,* E. A. Fox, ed., Blacksburg, VA: Department of Computer Science, Virginia Tech University, TR 93-95.

19. Fox, E. A., R. M. Akscyn, R. K. Furuta, and J. J. Leggett, "Digital Libraries." *Communications of the ACM, April 1995, Vol. 8, No. 4.*

20. French, J. C., E. A. Fox, K. Maly, and A. L. Selman, "Wide Area Technical Report Service: Technical Reports Online," *Communications of the ACM, April 1995, Vol. 8, No. 4.*

21. Lagoze, C., and J. R. Davis, "Dienst: An Architecture for Distrubuted Document Libraries," *Communications of the ACM, April 1995, Vol. 8, No. 4.*

22. *Proceedings of Digital Libraries '94, the First Annual Conference on the Theory and Practice of Digital Libraries,* College Station, Texas June 19-21, 1994.

23. Manjunath, B. S., "Image Browsing on the Alexandria Digital Library (ADL) Project," *d-Lib Magazine*, August 1995.

24. *Information Systems Research and Technology Applications for Navy Libraries,* memorandum from Joan Buntzen, Librarian of the Navy, to Dr. Ted Lewis, Chair Computer Science Department, NPS, 9 March 1995.

25. *Realizing the Information Future: The Internet and Beyond,* Washington, DC: National Academy Press, 1994.

26. *The University of Michigan Digital Library Project,* February 1994.

27. *Proposal to Re-Engineer Scientific-Technical Information (STI) Networking in the Navy Laboratory/Center (NLC) Community*, March-April 1995.

28. Christel, T., T. Kanade, M. Mauldin, R. Reddy, M. Sirbu, S. Stevens, and H. Wactlar, *Informedia Digital Video Library, (http://fuzine.mt.cs.cmu.edu/im/informedia.html).*

29. The Standford Digital Libraries Group, *The Stanford Digital Library Project, (http://www-diglib.stanford.edu/diglib).*

30. Wilensky, Robert, *UC Berkeley's Digital Library Project, (http://http.cs.berkeley.edu~wilensky).*

31. Smith, T. R., and J. Frew, Alexandria Digital Library, (http://alexandria.sdc.ucsb.edu/).

32. Schatz, Bruce, *Building the Interspace: The Illinois Digital Library Project, (http://www. grainger.uiuc.edu/dli).*

33. Crum, Laurie, *University of Michigan Digital Library Project, (http://www.sils.umich.edu/UMDL/ HomePage.html).*

34. Purday, Jonathan, "The British Library's Initiatives for Access Project," *Communications of the ACM, April 1995, Vol. 8, No. 4.*

35. Becker, Herbert S., "Library of Congress Digital Library Effort," *Communications of the ACM, April 1995, Vol. 8, No. 4.*

36. Wiederhold, Gio, "*Digital Libraries,* Value and Productivity," *Communications of the ACM, April 1995, Vol. 8, No. 4.*

37. *Wide Area Technical Report Service (WATERS), (http://www.cs.odu.edu/WATERS/WATERS-GS.html).*

38. Rao, R., J., Pedersen, M. Hearst, and J. Mackinlay, "Rich Interaction in the Digital Library," *Communications of the ACM, April 1995, Vol. 8, No. 4.*

39. *What is the NII?, (http://www.niiac.info.org/~niiac/main_doc.html).*

40. Farber, David, "Living in the Global Information Infrastructure -- the potentials and concerns," *The Alfred Fitler Moore Professor of Telecommunication Systems*, University of Pennsylvania.

41. Marshall, Catherine, P. J. Nurnberg, R. Furuta, J. Leggett, and F. Shipman III, "Digital Libraries: Issues and Architectures," *Proceeding of Digital Libraries '95 Conference,* (Austin, TX, June 11-13), *(http://www.csdl. tamu. edu/DL95/)*.

42. *Intelligent Information Access, (http://www.parc.xerox.com).*

43. Griffiths, Jose-Marie, K. K. Kertis, "Access to Large Digital Libraries of Scientific Information Across Networks," *Proceeding of Digital Libraries '95 Conference,* (Austin, TX, June 11-13), *(http://www.csdl.tamu.edu/DL95/)*.

44. Sullivan, G. R. (Gen), J. M. Dubik (Col.), *Army Digitization Master Plan, (http://www.ado.army.mil).*

45. Berwick, R. C., J. M. Carroll, *Research Priorities for the World-Wide Web, (http://www.cc.gatech.edu/gvu/nsf-ws/report/Report.html).*

46. Kroenke, David M., *Database Processing,* Englewood Cliffs, NJ: Prentice Hall, 1995, pp. 29-30.

47. *An Overview of the INQUERY System as Used for the TIPSTER Project, (http://ciir.cs.umass.edu/ info/inquery.html).*

48. *Introduction to SGML, (http://www.auto-graphics.com/www/publishing/intrsgml.html).*

49. *DB Connect Socrates, (http://lindy.stanford.edu:80/~asc/zform. CGI? PSOC).*

50. *The Java Language Environment: A White Paper, (http://www.javasoft. com/whitePaper/java-whitepaper-1.html).*

51. *NSF Announces Awards for Digital Libraries Research, (http://www-diglib. standford.edu/diglib/pub/nsf.announce.html).*

52. Croft, Bruce W., "NSF Center for Intelligent Information Retrieval," *Communications of the ACM, April 1995, Vol. 8, No. 4.*

53. Rowe, Neil C., *Proposal for Research -- Reimbursable Funding*, ARPA/ CSTO, October 1995, pp. 3-5.

54. Department of Defense Technical Architecture Framework for Information Management, DISA,Vol. 1-8, November 1993.

55. The Report of the Working Group on Intellectual Property Rights: *Intellectual Property and the National Information Infrastructure*, Washington, DC, September 1995.

56. Bhargava, H. K., R. Krishman, R. Muller, *Decision Support on Demand: Emerging Electronic Markets for Decision Technologies*, February 1996.

57. Jones, Carl R., *ITM in DOD Course,* Class Notes, Naval Postgraduate School, Monterey, CA, September 1995.

58. Brutzman, D., *NPSDL Working Papers*, Naval Postgraduate School, Monterey, CA, August 1995.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center 2
   8725 John J. Kingman Road, Suite 0944
   Fort Belvoir, VA 22060-6218

2. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Road
   Monterey, CA 93943-5101

3. Director, Training and Education 1
   MCCDC (Code C46)
   1019 Elliot Road
   Quantico, VA 22134-5027

4. Larry Clemens 1
   United States Naval Academy
   Nimitz Library
   Annapolis, MD 21402-5029

5. Naval General Library Program Office 1
   Naval Education and Training Program Management Support Activity
   Pensacola, FL 32509-5100

6. Naval Institute 1
   Photographic and Reference Library
   118 Maryland Ave.
   Annapolis, MD 21402-5035

7. Roderick D. Atkinson 1
   Naval Research Library
   Ruth H. Hooker Research Library and Technical Info Center
   4555 Overlook Ave.
   SW, Washington, DC 20375-5000

8. Pearce S. Grove 1
Library Director
Marine Corps University
2040 Broadway St.
Quantico, VA 22134-5107

9. Marine Corps Technical Library 1
Research, Development and Acquisition Command (Code PSD)
Washington, DC 20380-0001

10. James T. Helling 1
Director
AFIT Library
2950 P Street Bldg 642
Wright-Patterson AFB, OH 45433

11. Naval Surface Warfare Officers School Library 1
CO, Surface Warfare Officers School Command
Newport, RI 02841-5012

12. Joan L. Buntzen 1
Librarian of the Navy
Navy Department Library
Bldg 44, Washington Navy Yard
Washington, DC 20374-0571

13. Bohdan Kohutiak 1
Director, Libraries
US Army War College
Carlisle Barracks
Carlisle, PA 17013-5050

14. Robert B. Lane 1
Dir., Air University Libraries
USAF Air University
600 Chennault Circle
Maxwell, AFB, AL 36112-2606

15. Sarah Mikel 1
Library Director
National Defense University
Bldg 62, Marshall Hall
Washington, DC 20319-6000

16. Robert E. Schnare 1
Director
U.S. Naval War College
686 Cushing Road
Newport, RI 02841-1207

17. Laurie Stackpole 1
Chief Librarian
Naval Research Laboratory Library
Code 5220
Washington, DC 20375-5334

18. Richard Werking 1
Library Director
US Naval Academy
Annapolis, MD 21402

19. Professor Thomas Wu (Code CS/Wq) 1
Naval Postgraduate School
Monterey, CA 93943-5100

20. Dr. Richard S. Elster, Provost 1
Naval Postgraduate School
Monterey, CA 93943-5101

21. Dr. Ted Lewis, Chair (Code CS) 1
Naval Postgraduate School
Monterey, CA 93943-5101

22. Dave Norman (Code 51). 1
Director, W.R. Church Computer Center
Naval Postgraduate School
Monterey, CA 93943-5101

23. Dr. Craig Rasmussen (Code MA/Ra) 1
Naval Postgraduate School
Monterey, CA 93943-5101

24. Dr. Maxine Reneker (Code 52) 1
Director, Dudley Knox Library
Naval Postgraduate School
Monterey, CA 93943-5101

25. CAPT George Zolla, USN (Code 05b) 1
Naval Postgraduate School
Monterey, CA 93943-5101

26. Dr. Arthur St. George 1
National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230

27. Dr. Suresh Sridhar (Code SM/Sr) 1
Naval Postgraduate School
Monterey, CA 93943-5103

28. Chris Taylor 1
Director of Computing and Computer Resources (CCR)
California State University, Monterey Bay
100 Campus Center
Seaside, CA 93955-8001

29. Steve Watkins 1
Director of Library Services
California State University, Monterey Bay
100 Campus Center
Seaside, CA 93955-8001

30. Honorable Henry Mello 1
1200 Aguajito Road
Monterey, CA 93940

31. Doug Gould (Code 52) 1
Naval Postgraduate School
Monterey, CA 93943-5101

32. Prof. Hemant Bhargava (Code SM/Bh) 15
Naval Postgraduate School
Monterey, CA 93943-5103

33. Prof. Don Brutzman (Code UW/Br) 1
Naval Postgraduate School
Monterey, CA 93943-5100

34. Ellen Pastore 1
MOBAC
c/o: MPC Library Bldg
980 Fremont St.
Monterey, CA 93940

35. Colonel William Norris 1
1121 Sagebrush TR SE
Alberquerque, NM 87123

36. Robert E. Norris 2
810 Pine Ave.
Pacific Grove, CA 93950

37. David S. West 2
3820 Bergin Drive
Monterey, CA 93940